```
Db       1316 isLeuLeuLysGluLysAsnGlu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArg 1335

Qy       3769 ATT---CAATATGAAAAAGAGAAAGAGAAACAGAAACTCATG 3807
              |||   |||||||||||||||||||   :::   :::::::::
Db       1336 IleTyrGlnTyrGluLysGluLysAlaGluThrGluValIle 1349

RESULT 9
US-10-007-805-553
; Sequence 553, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 553
;   LENGTH: 1013
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-007-805-553

Alignment Scores:
Pred. No.:                4.54e-80        Length:          1013
Score:                    926.50          Matches:          504
Percent Similarity:       59.39%          Conservative:     116
Best Local Similarity:    48.28%          Mismatches:       293
Query Match:              13.08%          Indels:           144
DB:                       13              Gaps:              83

US-09-602-362E-22 (1-4115) x US-10-007-805-553 (1-1013)

Qy       1018 CACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAATTACGTG---GGCAGAAAA 1074
              |||:::|||                |||||||||          :::             |||
Db          8 HisValGlySerMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLys 27

Qy       1075 GGAAGACCTAGGAAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT- 1129
              |||||||||||||||   |||||||||||||||:::   |||   |||   ||||||
Db         28 GlyArgProArgLysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysVal 47

Qy       1130 GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCA 1185
              |||||||||||||||||||   |||         |||   :::|||||||||   |||
Db         48 AlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAla 67

Qy       1186 TGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAG 1239
              ||||||   |||   :::      ::::::      :::|||||||||         :::
Db         68 CysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGlu 87

Qy       1240 ATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAG 1292
                :::      |||||||||||||||||   |||||||||   ||||||
Db         88 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSer 106

Qy       1293 TTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAAT 1349
                |||||||||||   :::      |||:::|||   |||      |||||||||
Db        107 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 126

Qy       1350 AAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA 1406
              ::::::      :::         |||||||||||||||||||||   ::::::   :::
Db        127 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 146

Qy       1407 ---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCA 1454
                        |||||||||                |||||:::          :::   |||
Db         147 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAla 164

Qy       1455 GATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCT 1510
              ||||||   |||||||||   :::      |||          |||||||||||||||||
Db         165 AspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer 184

Qy       1511 AGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCA 1564
              |||||||||   |||||||||   ::::::   :::          ||||||||||||||
Db         185 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGl 204

Qy       1565 AAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTC 1618
              |||||||       |||   |||||||||      :::||||||        |||||||
Db         204 nLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuL 224

Qy       1619 TGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGAC 1672
              ||||||||||           ::::::|||||||||   |||          :::::|||
Db         224 euLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAsp 243

Qy       1673 ATGCAAACTTTCAAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAA 1728
              |||||||||||||||   ||||||||||                        :::||||||
Db         244 MetGlnThrPheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGlu 263

Qy       1729 ATC---AAAAGTCTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGA 1785
              :::   ||||||:::         :::|||   :::|||||||||:::
Db         264 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 283

Qy       1786 TGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACT 1839
                 |||   |||||||||||||||||||:::          :::      |||||||||
Db         284 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspThr 303

Qy       1840 GGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTC 1892
              |||||||||   |||||||||||||||||||   |||||||||:::
Db         304 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 323

Qy       1893 AAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCT 1949
              |||   |||||||||||          :::|||   |||          |||||||||
Db         323 nLysGluIleAspLysIleAsnGlyLysLeuGluGlySerProGlyLysAspGlyLeuLe 343

Qy       1950 GAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACA 2003
              |:::          |||   :::   |||||||||||||||||        :::   ||||
Db         343 uLysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspM 363

Qy       2004 TGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA- 2059
              ||||||||   |||   :::         :::|||||||||||||||||||   ::::::
Db         363 etGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluM 383

Qy       2060 --TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAG 2114
                  |||         ||||||         |||   :::|||||||||||||||||||
Db         383 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 402

Qy       2115 ATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTC 2168
                 |||   |||||||||   :::      |||          |||||||||||||||||
Db         403 Asp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSe 422

Qy       2169 TGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATC 2222
              | |||||||||       |||||||||::::::   |||          :::|||   ||||
Db         422 r-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisG 442

Qy       2223 AAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCT 2278
              |||||||       |||         |||:::|||||||||   ||||||:::   |||
Db         442 lnLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheL 462

Qy       2279 GAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCA 2332
                 :::   |||       :::   |||||||||   |||      :::      |
Db         462 euLysSerProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspM 482

Qy       2333 TGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT 2389
              |||||||||||||||||       |||   :::          |||||||||||:::
Db         482 etGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-Glu 501

Qy       2390 ---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAG 2443
```

```
Db        502 MetGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuAr 520

Qy       2444 ATCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGA 2494

Db        520 gAlaAspGlnMetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAs 540

Qy       2495 TTCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACA 2548

Db        540 pSerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThr 559

Qy       2549 CACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT 2605

Db        560 HisGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLys 579

Qy       2606 ---CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAA 2659

Db        580 IleLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGlu 599

Qy       2660 CA-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC--- 2712

Db        600 GlnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeu 619

Qy       2713 TGTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGA 2769

Db        620 SerGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGl 639

Qy       2770 G---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATA 2823

Db        639 uLeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspI 659

Qy       2824 TATTAAAGAAAAAA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAG 2871

Db        659 leLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------ 676

Qy       2872 AAAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAAT 2925

Db        677 LysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLe 696

Qy       2926 TGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTA 2979

Db        696 uLysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrL 716

Qy       2980 TTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCA 3036

Db        716 euLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaT 736

Qy       3037 CACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAA 3092

Db        736 hrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeu 755

Qy       3093 AGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG 3148

Db        756 LysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLys 775

Qy       3149 ---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATT 3202

Db        776 ArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuT 795

Qy       3203 CTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACC 3255

Db        795 hrSerLysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisH 815

Qy       3256 ATCCTGACTGGCTTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAG 3303

Db        815 isPro---------ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerA 832

Qy       3304 AAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGA 3357

Db        832 rgLysSerGlnGluProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetA 852

Qy       3358 ATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGA 3411

Db        852 snValAspValSerSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGlu 871

Qy       3412 GCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAG 3468

Db        872 AlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuAr 891

Qy       3469 A---AATACATTGGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT- 3521

Db        891 gGluAsnThrLeuValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnM 911

Qy       3522 --GAAGGAAGCGAACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTG 3573

Db        911 etLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluG 931

Qy       3574 AACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAG 3633

Db        931 lnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGln 950

Qy       3634 CAATT----AGTCATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------- 3680

Db        951 GlnGlnLeuValHisAlaHisLys-----------LysAlaAspAsnLysSerLysIle 966

Qy       3681 ------GATATTCATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAG 3731

Db        967 ThrIleAspIleHisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsn 986

Qy       3732 AGGAGATATT---AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGA 3785

Db        987 Glu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGl 1006

Qy       3786 GAAAGAGAAACAGAAACTCATG 3807

Db       1006 uLysAlaGluThrGluValIle 1013
```

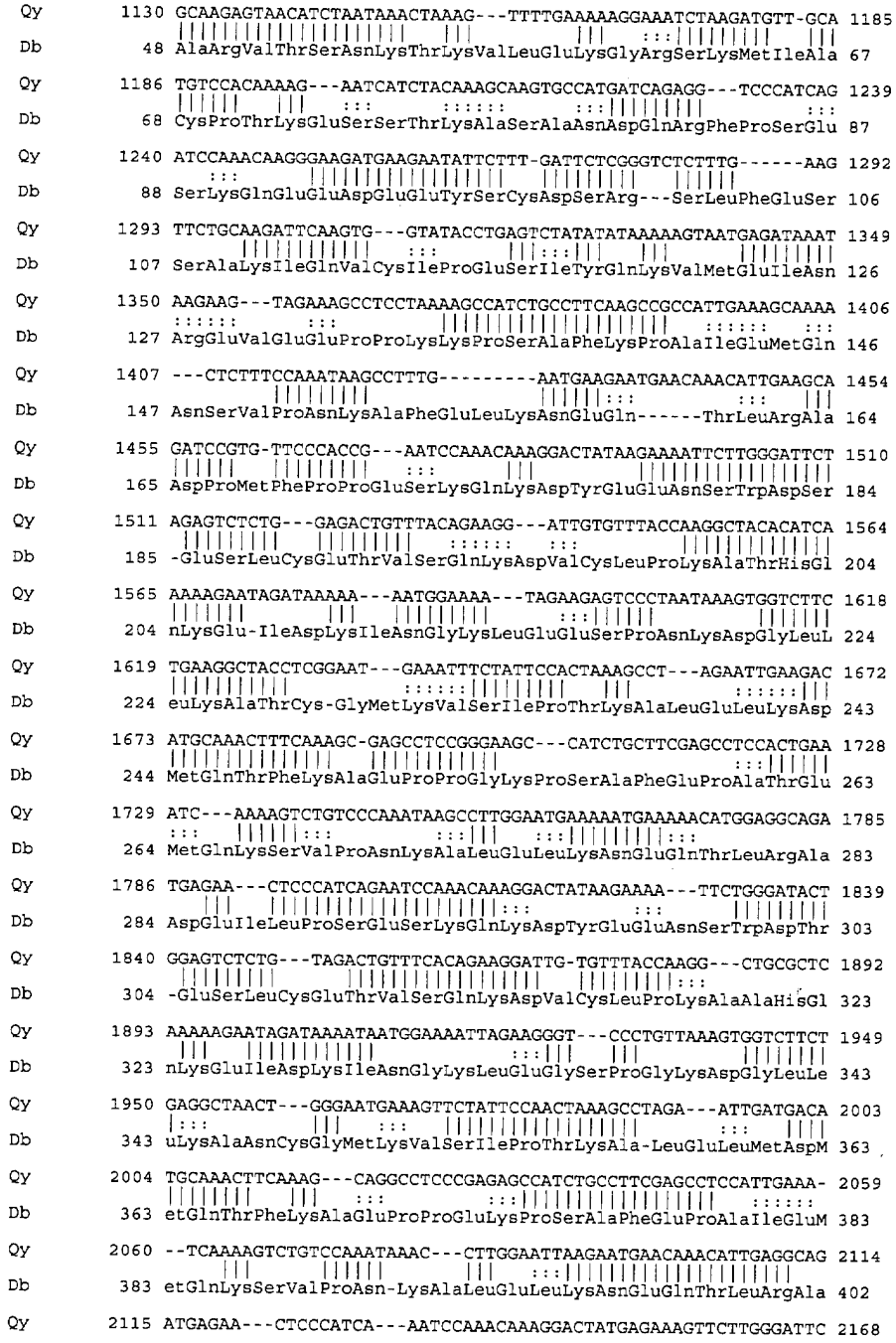
RESULT 10
US-10-076-622-553
; Sequence 553, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 553
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-553

Alignment Scores:

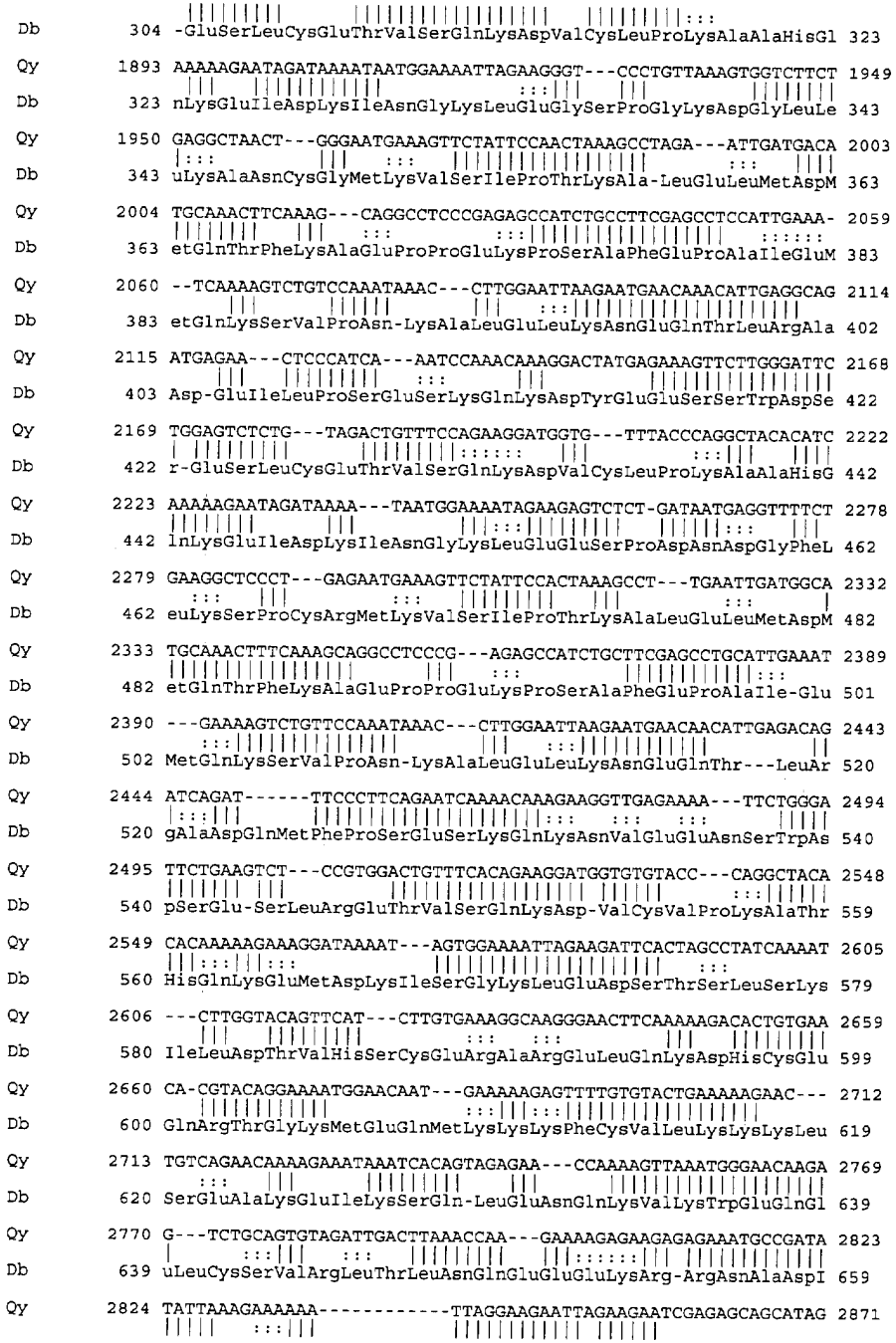
Pred. No.:	4.54e-80	Length:	1013
Score:	926.50	Matches:	504
Percent Similarity:	59.39%	Conservative:	116
Best Local Similarity:	48.28%	Mismatches:	293
Query Match:	13.08%	Indels:	144
DB:	14	Gaps:	83

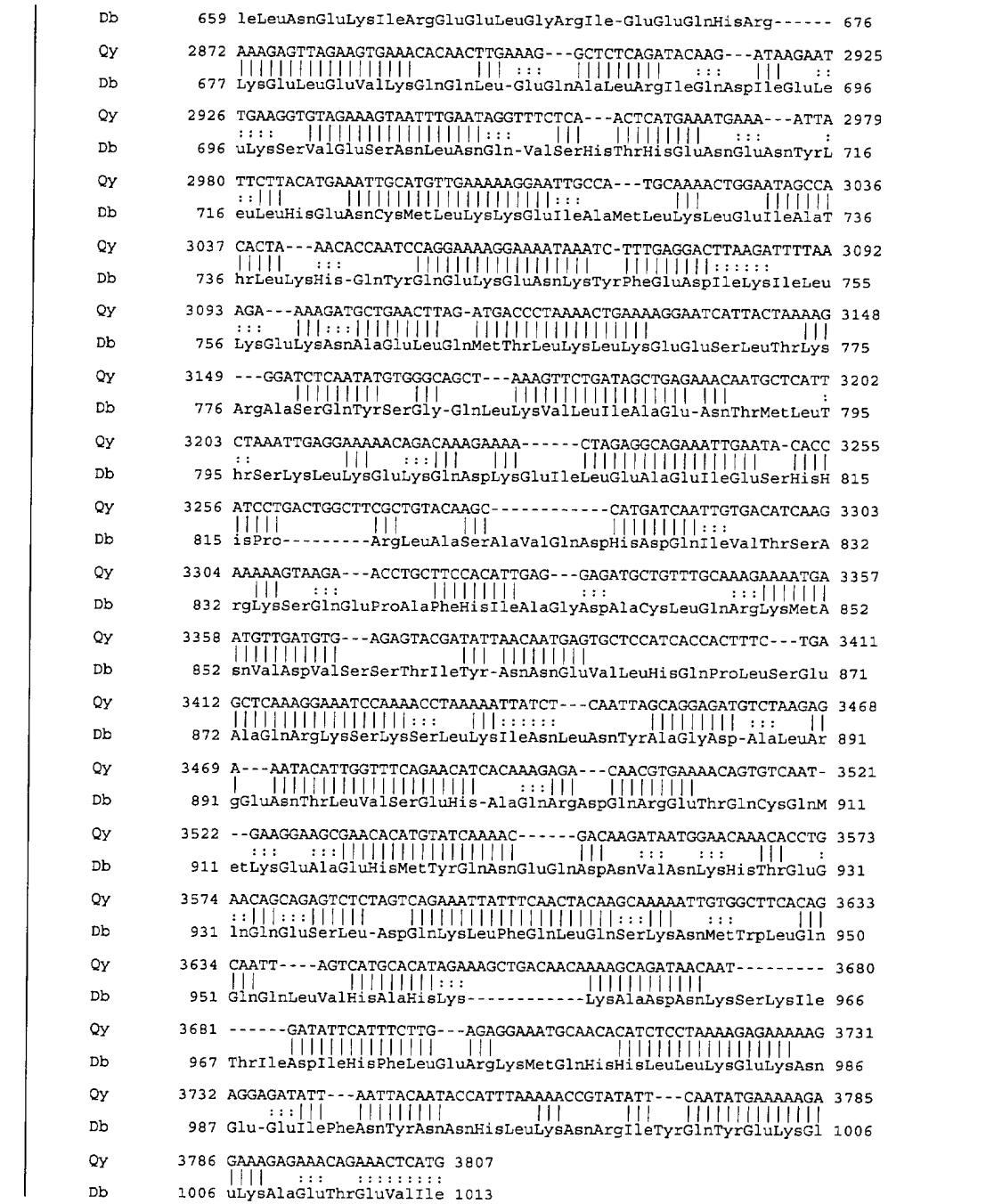
US-09-602-362E-22 (1-4115) x US-10-076-622-553 (1-1013)

```
Qy       1018 CACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAATTACGTG---GGCAGAAAA 1074

Db          8 HisValGlySerMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLys 27

Qy       1075 GGAAGACCTAGGAAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT- 1129

Db         28 GlyArgProArgLysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysVal 47
```


Qy 1130 GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCA 1185
Db 48 AlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAla 67
Qy 1186 TGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAG 1239
Db 68 CysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGlu 87
Qy 1240 ATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAG 1292
Db 88 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSer 106
Qy 1293 TTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAAT 1349
Db 107 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 126
Qy 1350 AAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA 1406
Db 127 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 146
Qy 1407 ---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCA 1454
Db 147 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAla 164
Qy 1455 GATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCT 1510
Db 165 AspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer 184
Qy 1511 AGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCA 1564
Db 185 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGl 204
Qy 1565 AAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTC 1618
Db 204 nLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuL 224
Qy 1619 TGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGAC 1672
Db 224 euLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAsp 243
Qy 1673 ATGCAAACTTTCAAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAA 1728
Db 244 MetGlnThrPheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGlu 263
Qy 1729 ATC---AAAAGTCTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGA 1785
Db 264 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 283
Qy 1786 TGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACT 1839
Db 284 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspThr 303
Qy 1840 GGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTC 1892
Db 304 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 323
Qy 1893 AAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCT 1949
Db 323 nLysGluIleAspLysIleAsnGlyLysLeuGluGlySerProGlyLysAspGlyLeuLe 343
Qy 1950 GAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACA 2003
Db 343 uLysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspM 363
Qy 2004 TGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA- 2059
Db 363 etGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluM 383
Qy 2060 --TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAG 2114
Db 383 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 402
Qy 2115 ATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTC 2168
Db 403 Asp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSe 422
Qy 2169 TGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATC 2222
Db 422 r-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisG 442
Qy 2223 AAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCT 2278
Db 442 lnLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheL 462
Qy 2279 GAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCA 2332
Db 462 euLysSerProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspM 482
Qy 2333 TGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT 2389
Db 482 etGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-Glu 501
Qy 2390 ---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAG 2443
Db 502 MetGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuAr 520
Qy 2444 ATCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGA 2494
Db 520 gAlaAspGlnMetProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAs 540
Qy 2495 TTCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACA 2548
Db 540 pSerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThr 559
Qy 2549 CACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT 2605
Db 560 HisGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLys 579
Qy 2606 ---CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAA 2659
Db 580 IleLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGlu 599
Qy 2660 CA-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC--- 2712
Db 600 GlnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeu 619
Qy 2713 TGTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGA 2769
Db 620 SerGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGl 639
Qy 2770 G---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATA 2823
Db 639 uLeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspI 659
Qy 2824 TATTAAAGAAAAAA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAG 2871
Db 659 leLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------ 676
Qy 2872 AAAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAAT 2925
Db 677 LysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLe 696
Qy 2926 TGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTA 2979
Db 696 uLysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrL 716
Qy 2980 TTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCA 3036
Db 716 euLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaT 736
Qy 3037 CACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAA 3092
Db 736 hrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeu 755
Qy 3093 AGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG 3148

```
Db        756 LysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLys 775

Qy       3149 ---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATT 3202
                  |||||||||  |||     |||||||||||||||||||  |||         :
Db        776 ArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuT 795

Qy       3203 CTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACC 3255
                ::        |||   :::|||   |||      ||||||||||||||||||   ||||
Db        795 hrSerLysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisH 815

Qy       3256 ATCCTGACTGGCTTCGCTGTACAAGC-----------CATGATCAATTGTGACATCAAG 3303
                |||||          |||      |||           |||||||||:::
Db        815 isPro---------ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerA 832

Qy       3304 AAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGA 3357
                  |||   :::        |||||||||      :::            :::|||||||
Db        832 rgLysSerGlnGluProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetA 852

Qy       3358 ATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGA 3411
                |||||||||||           ||| |||||||||
Db        852 snValAspValSerSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGlu 871

Qy       3412 GCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAG 3468
                |||||||||||||||||:::   |||::::::          |||||||| :::   ||
Db        872 AlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuAr 891

Qy       3469 A---AATACATTGGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT- 3521
                |   ||||||||||||||||||||    :::|||   |||||||||
Db        891 gGluAsnThrLeuValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnM 911

Qy       3522 --GAAGGAAGCGAACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTG 3573
                  :::   :::|||||||||||||||||||      |||   :::   :::   |||  :
Db        911 etLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluG 931

Qy       3574 AACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAG 3633
                ::|||:::||||||    |||||||||||||||||||||:::|||   :::      |||
Db        931 lnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGln 950

Qy       3634 CAATT----AGTCATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------- 3680
                |||          |||||||||:::            ||||||||||
Db        951 GlnGlnLeuValHisAlaHisLys-----------LysAlaAspAsnLysSerLysIle 966

Qy       3681 ------GATATTCATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAG 3731
                     ||||||||||||||   |||           |||||||||||||||||
Db        967 ThrIleAspIleHisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsn 986

Qy       3732 AGGAGATATT---AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGA 3785
                  :::|||   |||||||||          |||      |||   ||||||||||||
Db        987 Glu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGl 1006

Qy       3786 GAAAGAGAAACAGAAACTCATG 3807
                ||||   :::   :::::::::
Db       1006 uLysAlaGluThrGluValIle 1013

RESULT 11
US-10-124-805-553
; Sequence 553, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C12
;  CURRENT APPLICATION NUMBER: US/10/124,805
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 553
;   LENGTH: 1013
;   TYPE: PRT
;   ORGANISM: Homo sapiens

US-10-124-805-553

Alignment Scores:
Pred. No.:                4.54e-80      Length:          1013
Score:                    926.50        Matches:         504
Percent Similarity:       59.39%        Conservative:    116
Best Local Similarity:    48.28%        Mismatches:      293
Query Match:              13.08%        Indels:          144
DB:                       14            Gaps:            83

US-09-602-362E-22 (1-4115) x US-10-124-805-553 (1-1013)

Qy       1018 CACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAATTACGTG---GGCAGAAAA 1074
                |||:::|||                |||||||||          :::             |||
Db          8 HisValGlySerMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLys 27

Qy       1075 GGAAGACCTAGGAAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT- 1129
                |||||||||||||||   |||||||||||||||:::   |||   |||   ||||||
Db         28 GlyArgProArgLysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysVal 47

Qy       1130 GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCA 1185
                |||||||||||||||||||  |||      |||  :::|||||||||  |||
Db         48 AlaArgValThrSerAsnLysThrLysGluLysGlyArgSerLysMetIleAla 67

Qy       1186 TGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAG 1239
                ||||||   |||   :::      ::::::      :::|||||||||         :::
Db         68 CysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGlu 87

Qy       1240 ATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAG 1292
                  :::      |||||||||||||||||   |||||||||   ||||||
Db         88 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSer 106

Qy       1293 TTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAAT 1349
                      ||||||||||||   :::      |||:::|||   |||      |||||||||
Db        107 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 126

Qy       1350 AAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA 1406
                ::::::      :::          ||||||||||||||||||||   ::::::   :::
Db        127 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 146

Qy       1407 ---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCA 1454
                     |||||||||          ||||||:::      :::      |||
Db        147 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAla 164

Qy       1455 GATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCT 1510
                |||||   |||||||||   :::      |||          ||||||||||||||||
Db        165 AspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer 184

Qy       1511 AGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCA 1564
                 |||||||||   |||||||||   ::::::   :::          ||||||||||||
Db        185 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGl 204

Qy       1565 AAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTC 1618
                |||||||          |||  |||||||||      :::||||||          ||||||
Db        204 nLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuL 224

Qy       1619 TGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGAC 1672
                |||||||||||          ::::::|||||||||   |||         ::::::|||
Db        224 euLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAsp 243

Qy       1673 ATGCAAACTTTCAAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAA 1728
                |||||||||||||||   ||||||||||||                  :::||||||
Db        244 MetGlnThrPheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGlu 263

Qy       1729 ATC---AAAAGTCTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGA 1785
                :::   |||||:::         :::|||   :::||||||||:::
Db        264 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 283

Qy       1786 TGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACT 1839
                  |||   |||||||||||||||||:::         :::      |||||||||
Db        284 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspThr 303

Qy       1840 GGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTC 1892
```





RESULT 12
US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

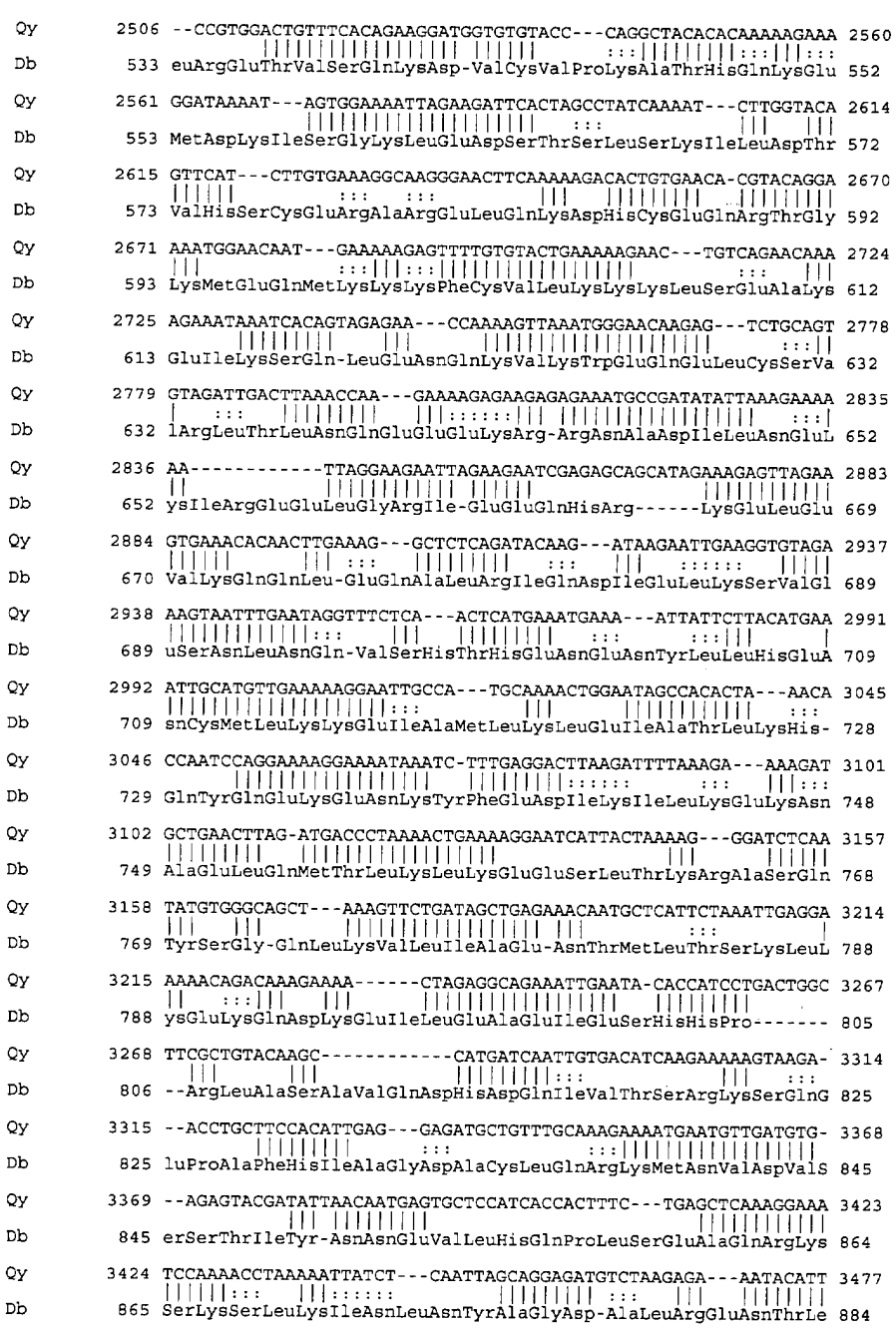
Alignment Scores:

Pred. No.:	3.76e-79	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	9	Gaps:	83

US-09-602-362E-22 (1-4115) x US-09-604-287A-475 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
```

```
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533
```

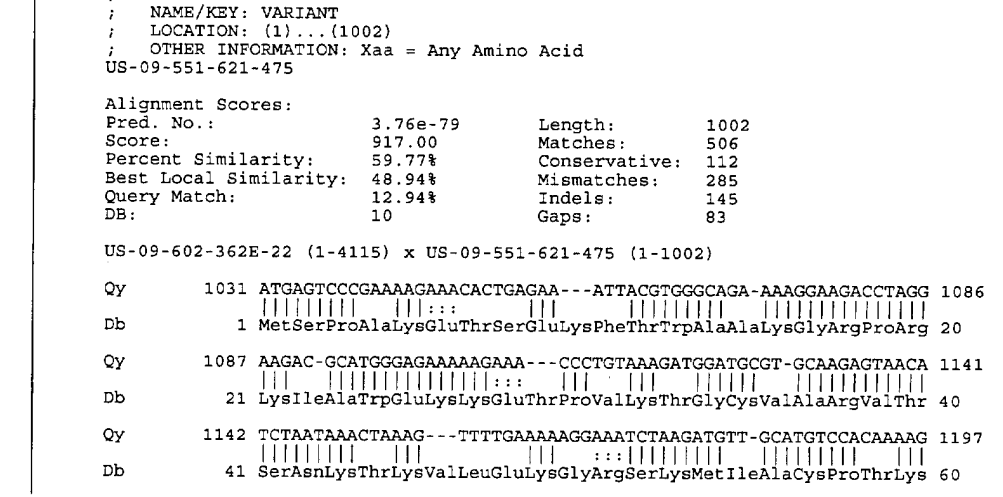


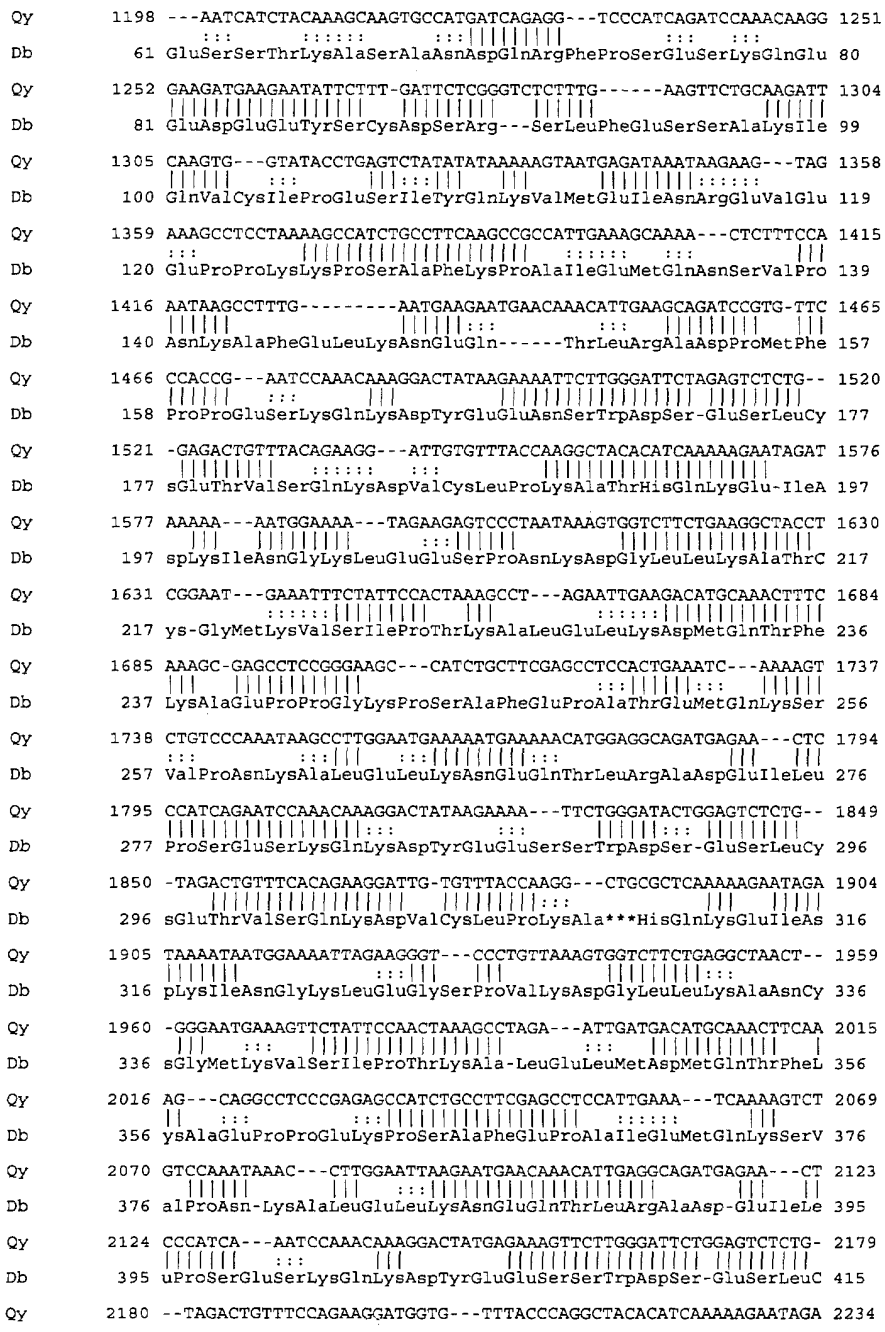
RESULT 13
US-09-551-621-475
; Sequence 475, Application US/09551621
; Publication No. US20030104366A1
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Alignment Scores:

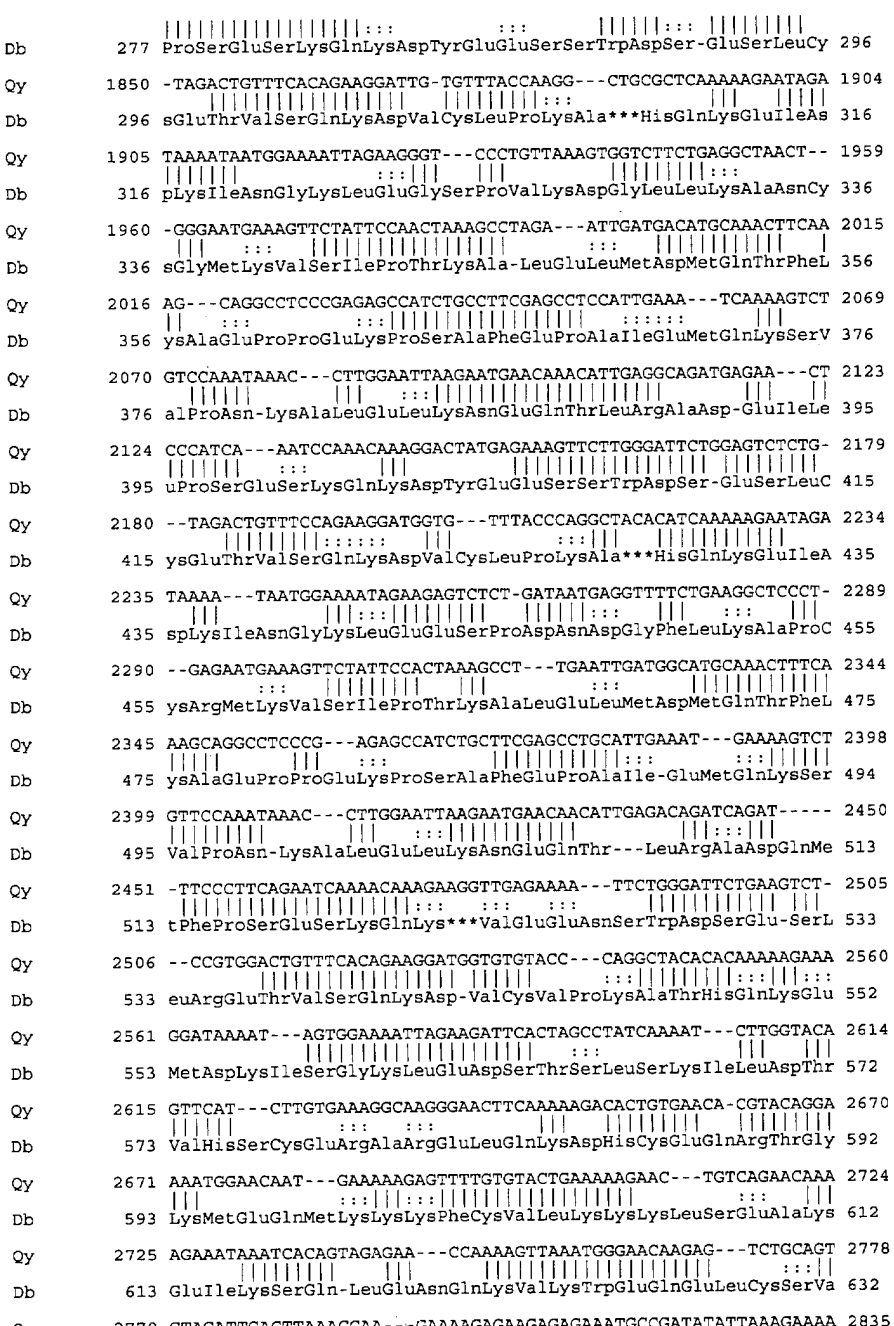
Pred. No.:	3.76e-79	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	10	Gaps:	83

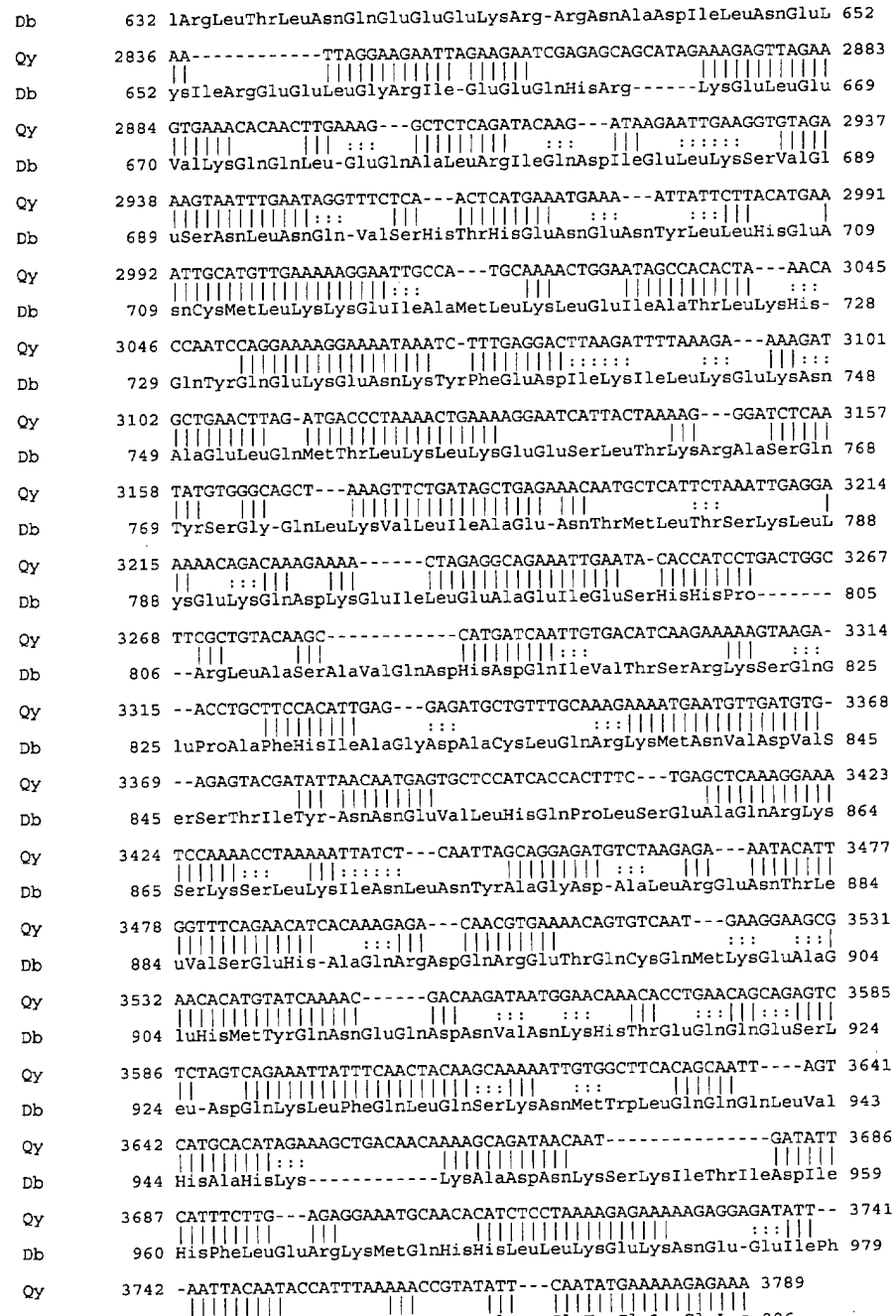
US-09-602-362E-22 (1-4115) x US-09-551-621-475 (1-1002)




Qy 1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db 61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGluSerLysGlnGlu 80
Qy 1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db 81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99
Qy 1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db 100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119
Qy 1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db 120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139
Qy 1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db 140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157
Qy 1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db 158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177
Qy 1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db 177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197
Qy 1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db 197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217
Qy 1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db 217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236
Qy 1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db 237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256
Qy 1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db 257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276
Qy 1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db 277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296
Qy 1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db 296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316
Qy 1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db 316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336
Qy 1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db 336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356
Qy 2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db 356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376
Qy 2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db 376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395
Qy 2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db 395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415
Qy 2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db 415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435
Qy 2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db 435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455
Qy 2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db 455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475
Qy 2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db 475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494
Qy 2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db 495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513
Qy 2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db 513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533
Qy 2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db 533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552
Qy 2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db 553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572
Qy 2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db 573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592
Qy 2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db 593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612
Qy 2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
Db 613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 632
Qy 2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
Db 632 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 652
Qy 2836 AA----------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
Db 652 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 669
Qy 2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
Db 670 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 689
Qy 2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
Db 689 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 709
Qy 2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
Db 709 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 728
Qy 3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
Db 729 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 748
Qy 3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
Db 749 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 768
Qy 3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214

```
Db        769 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 788

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267

Db        788 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 805

Qy       3268 TTCGCTGTACAAGC----------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314

Db        806 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 825

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368

Db        825 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 845

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423

Db        845 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 864

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477

Db        865 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 884

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531

Db        884 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 904

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585

Db        904 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 924

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641

Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686

Db        944 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741

Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789

Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996

RESULT 14
US-10-007-805-475
; Sequence 475, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 475
;   LENGTH: 1002
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 310, 429, 522
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-10-007-805-475

Alignment Scores:
Pred. No.:              3.76e-79      Length:          1002
Score:                  917.00        Matches:          506
Percent Similarity:     59.77%        Conservative:     112
Best Local Similarity:  48.94%        Mismatches:       285
Query Match:            12.94%        Indels:           145
DB:                     13            Gaps:              83

US-09-602-362E-22 (1-4115) x US-10-007-805-475 (1-1002)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086

Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141

Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197

Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251

Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304

Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358

Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415

Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465

Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520

Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576

Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630

Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684

Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737

Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794

Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
```





RESULT 15
US-10-076-622-475
; Sequence 475, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-076-622-475

Alignment Scores:

Pred. No.:	3.76e-79	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	14	Gaps:	83

US-09-602-362E-22 (1-4115) x US-10-076-622-475 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
```

```
Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
              |||||||||||||||||||||||   :::           |||   |||
Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
              ||||||         :::   :::         |||   |||||||||   |||||||||
Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
              |||           :::|||:::|||||||||||||||||||         :::   |||
Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
              |||||||||     |||       |||||||||||||||||||||       :::||
Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 632

Qy       2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
              |   :::   |||||||||   |||::::::||| ||||||||||||||||||   :::|
Db        632 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 652

Qy       2836 AA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
              ||            |||||||||||| ||||||                |||||||||||
Db        652 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 669

Qy       2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
              ||||||      ||| :::   |||||||||   :::   |||   ::::::   |||||
Db        670 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 689

Qy       2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
              |||||||||||||:::    |||   |||||||||   :::      :::|||     |
Db        689 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 709

Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
              ||||||||||||||||||||:::         |||      ||||||||||||   :::
Db        709 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 728

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
              |||||||||||||||||   |||||||||::::::      :::   |||:::
Db        729 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 748

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
              |||||||||   |||||||||||||||||||                |||      ||||||
Db        749 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 768

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
              |||   |||       ||||||||||||||||| |||         :::          |
Db        769 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 788

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
              ||   :::|||   |||       ||||||||||||||||||   ||||||||
Db        788 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 805

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
                |||       |||           |||||||||:::           |||   :::
Db        806 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 825

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
                |||||||||      :::                :::|||||||||||||||||||
Db        825 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 845

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
                  ||| |||||||||                           ||||||||||||
Db        845 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 864

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
              ||||||:::   |||::::::         |||||||||:::   |||   ||||||||
Db        865 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 884

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
              |||||||||||||    :::|||   |||||||||                 :::   :::|
Db        884 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 904

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
              |||||||||||||||||      |||   :::   :::   |||   :::|||:::||||
Db        904 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 924

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
              ||    |||||||||||||||||||:::|||   :::       ||||||
Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
              |||||||||:::           ||||||||||||                  ||||||
Db        944 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
              |||||||||   |||           ||||||||||||||||||       :::|||
Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
               |||||||||         |||      |||   ||||||||||||||||
Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996
```

Search completed: July 15, 2004, 09:44:59
Job time : 325.589 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 08:36:32 ; Search time 177.082 Seconds
(without alignments)
13131.604 Million cell updates/sec

Title: US-09-602-362E-22
Perfect score: 7086
Sequence: 1 ctagtctatacagcaacgac.....gagtggactccacctggaaa 4115

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093625_21981/app_query.fasta_1.10325
-DB=A_Geneseq_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09602362_@CGN_1_1_729_@runat_15072004_093625_21981 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1234	17.4	1341	4	AAB84702	Aab84702 Amino aci
2	1234	17.4	1341	5	ABJ05537	Abj05537 Breast ca
3	1234	17.4	1341	6	ABJ37784	Abj37784 Human tum
4	1234	17.4	1341	6	ABR47548	Abr47548 Breast ca
5	1223	17.3	1349	6	ABJ37788	Abj37788 Human tum
6	926.5	13.1	1013	6	ABJ37783	Abj37783 Human tum
7	917	12.9	1002	4	AAU33351	Aau33351 Human bre
8	917	12.9	1002	5	ABG78918	Abg78918 Human bre
9	917	12.9	1002	6	ABJ37741	Abj37741 Human tum
10	917	12.9	1002	8	ADE44427	Ade44427 Human bre
11	914.5	12.9	1095	4	AAU33357	Aau33357 Human bre
12	914.5	12.9	1095	5	ABG78924	Abg78924 Human bre
13	914.5	12.9	1095	6	ABJ37747	Abj37747 Human tum
14	612.5	8.6	1239	6	ABJ37789	Abj37789 Human tum
15	597	8.4	661	6	ABJ37782	Abj37782 Human tum
16	583.5	8.2	650	4	AAB50263	Aab50263 Human bre
17	583.5	8.2	650	4	AAG65983	Aag65983 B726P spl
18	583.5	8.2	650	4	AAU33346	Aau33346 Human bre
19	583.5	8.2	650	5	ABG78913	Abg78913 Human bre
20	583.5	8.2	650	6	ABJ37736	Abj37736 Human tum
21	583.5	8.2	650	8	ADE44421	Ade44421 Human bre
22	582	8.2	743	4	AAU33358	Aau33358 Human bre
23	582	8.2	743	5	ABG78925	Abg78925 Human bre
24	582	8.2	743	6	ABJ37748	Abj37748 Human tum
25	544.5	7.7	1225	6	ABR47547	Abr47547 Breast ca
26	481	6.8	1011	4	AAB84703	Aab84703 Amino aci
27	465	6.6	512	4	AAB84701	Aab84701 Amino aci
28	421.5	5.9	445	4	AAB50249	Aab50249 Human bre
29	421.5	5.9	445	4	AAG65987	Aag65987 B726P spl
30	421.5	5.9	445	4	AAU33350	Aau33350 Human bre
31	421.5	5.9	445	5	ABG78917	Abg78917 Human bre
32	421.5	5.9	445	6	ABJ37740	Abj37740 Human tum
33	421.5	5.9	445	8	ADE44425	Ade44425 Human bre
34	416.5	5.9	432	4	AAB50244	Aab50244 Human bre
35	416.5	5.9	432	4	AAG65982	Aag65982 B726P thi
36	416.5	5.9	432	4	AAU33345	Aau33345 Human bre
37	416.5	5.9	432	5	AAO15779	Aao15779 Human bre
38	416.5	5.9	432	5	ABG78912	Abg78912 Human bre
39	416.5	5.9	432	6	ABJ37735	Abj37735 Human tum
40	416.5	5.9	432	8	ADE44133	Ade44133 Human bre
41	408.5	5.8	466	4	AAB50248	Aab50248 Human bre
42	408.5	5.8	466	4	AAG65986	Aag65986 B726P spl
43	408.5	5.8	466	4	AAU33349	Aau33349 Human bre
44	408.5	5.8	466	5	ABG78916	Abg78916 Human bre
45	408.5	5.8	466	6	ABJ37739	Abj37739 Human tum

ALIGNMENTS

RESULT 1
AAB84702
ID AAB84702 standard; protein; 1341 AA.
XX
AC AAB84702;
XX
DT 17-SEP-2001 (first entry)
XX
DE Amino acid sequence of a human cancer associated antigen.
XX
KW Cancer associated antigen; ING1; tumour suppressor; cancer; vaccine.
XX
OS Homo sapiens.
XX
PN WO200147959-A2.
XX
PD 05-JUL-2001.
XX
PF 29-NOV-2000; 2000WO-US042334.
XX
PR 30-NOV-1999; 99US-00451739.
PR 24-OCT-2000; 2000US-00602362.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
PA (CORR) CORNELL RES FOUND INC.
XX
PI Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;
XX
DR WPI; 2001-441706/47.
XX
PT Isolated cancer associated nucleic acid molecule identified by SEREX
PT (serological identification of antigens by recombinant expression
PT cloning) technique, useful in nucleic acid based therapies to treat

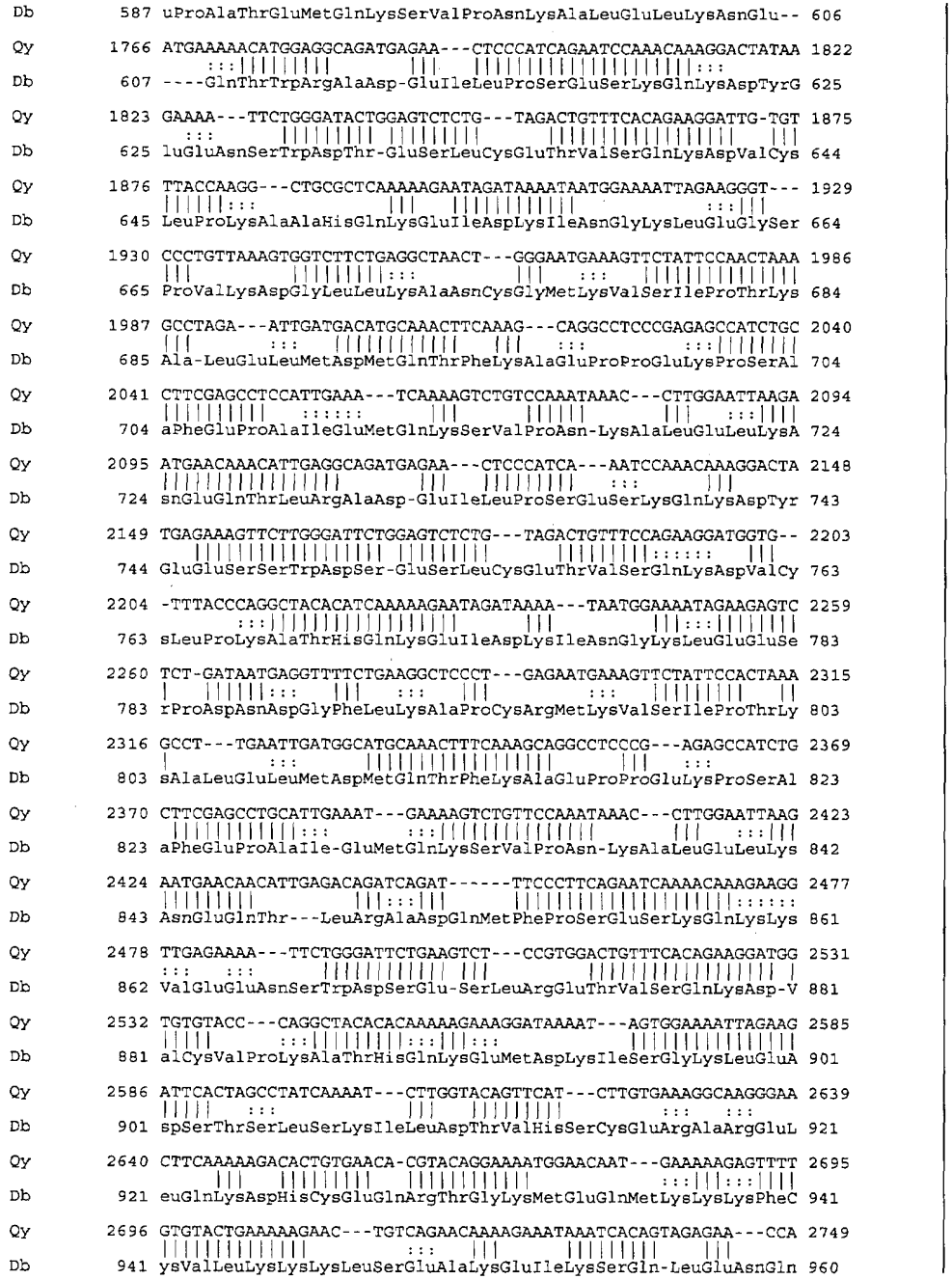
PT cancer.
XX
PS Claim 83; Page 53-57; 62pp; English.
XX
CC The present sequence represents a human cancer associated antigen. The
CC sequence was identified using probes derived from the ING1 gene. The ING1
CC gene is a tumour suppressor candidate gene. The cancer associated antigen
CC polynucleotides and polypeptides are useful for screening for the
CC possible presence of a pathological condition in a subject such as
CC cancer. The cancer associated antigen polypeptides are useful for
CC producing vaccines
XX
SQ Sequence 1341 AA;

Alignment Scores:

Pred. No.:	1.04e-131	Length:	1341
Score:	1234.00	Matches:	679
Percent Similarity:	59.18%	Conservative:	146
Best Local Similarity:	48.71%	Mismatches:	376
Query Match:	17.41%	Indels:	207
DB:	4	Gaps:	115

US-09-602-362E-22 (1-4115) x AAB84702 (1-1341)

```
Qy         96 CAAAGGGAAGAAGACATCAACCTTAATATACAAGAG---CCCAGAAGAGACTGCTCTAAC 152
Db          3 LysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArg-ThrAlaLeu-- 21

Qy        153 TGGGCCTGGTCAATGGCCT--------------GAGGAAGTAGTAACATTTC---GGTA 194
Db         22 ----------HisTrpAlaCysValAsnGlyHisGluGluValValThr-PheLeuValA 38

Qy        195 GACAGAAGTGCCAGCTGACGTCCTTG---TGGCGAACACGGACACCTCTGATGAAGGC-T 250
Db         38 spArgLysCysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaL 58

Qy        251 TACAATGCCACAGGAGGCTTTGC---AAATATTTGATAGATTC-GGTGCCGATATAAATC 306
Db         58 euGlnCysHisGlnGluAlaCysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnL 78

Qy        307 TCGTG---ATGTGTATGCAACATGGCTTCCATTATGCGTTTATAGTGGATTTTGTCAGT- 362
Db         78 euValAspValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerV 98

Qy        363 --GGTGGCAAACTGCTGTC-CATGGTGCAGTATCGAAGTGC---CAACAAGGCTGCCTCA 416
Db         98 alValAlaLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuT 118

Qy        417 CACCACTTTTACTACCAT---AACGAAAGAAGTGAGCAATTGTGGAATTTT---GCTGAT 470
Db        118 hrProLeuLeuLeu-SerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIle 137

Qy        471 AAAAATGCAAATGCGAATGCGTTA---ATAAGTTAAATGCACACCCTCATGCTGCTGTAT 527
Db        138 LysAsnAlaAsnAlaAsnAlaValAsnLys-TyrLysCysThrAlaLeuMetLeuAlaVa 157

Qy        528 G---TCTGGATCATCAGAGATAGTTGCATGCTTCTTAGCAAAA---TGTGACGTCTTT-- 579
Db        157 lCysHisGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAl 177

Qy        580 -GTGCAGATATAGTGGAGTAACTGCAGAACAT---ATGCTGTTACTGTGGATTTCTCACA 635
Db        177 aAlaAspIle-CysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisI 197

Qy        636 TTCATAACAAA---TTATGAATATATACGAAAATTATCAAAAATCA---TCAAATACCAA 689
Db        197 leHisGluGlnIleMet-GluTyrIleArgLysLeuSer-LysAsnHisGlnAsnThrAs 216

Qy        690 T---CAGAAGGAACTCTGCAGGAAACCTGATGAGGCTGCACCCT---GGCGGAAAGACAC 743
Db        216 nProGluGlyThrSerAlaGlyThrProAspGluAlaAla-ProLeuAlaGluArgThrP 236

Qy        744 CTGACACGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCACCCT---GGTGGA 797
Db        236 roAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAla-ProLeuValGlu 255

Qy        798 AAGACACCTGACAGGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCATCT--- 851
Db        256 ArgThrProAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeu 275

Qy        852 TGGTGGAGGAACATCTGACAAATTCAATGTTGGAGAAAGG---ACATCTGGAAAGTTCGA 908
Db        276 Val-GluGlyThrSerAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGl 295

Qy        909 ACGTCAGC---AGAAAAACACCTAGGAAATTACGATCC---TGCAAAAAAACATCTGAGA 962
Db        295 uGln-SerAlaGluGluThrProArgGluIleThrSerProAlaLysGluThrSerGluL 315

Qy        963 AATTTACG---GGCCAGCAAAGGAAGACCTAGAAGATCGCAGGGAGAA---AAAGAAGAC 1016
Db        315 ysPheThrTrpPro-AlaLysGlyArgProArgLysIleAlaTrpGluLysLysGluAsp 334

Qy       1017 ACACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAA--------------TTA 1061
Db        335 ThrProArgGluIleMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAla 354

Qy       1062 CGTGGGCAGAAAAGGAAGACCTAGGAAGACGCATGGGAGAAAAAGAAA---CCCTGTAAA 1118
Db        355 LysGlyArgProArgLysIle---------AlaTrpGluLysLysGluThrProValLys 371

Qy       1119 GATGGATGCGT-GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCT 1174
Db        372 ThrGlyCysValAlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSer 391

Qy       1175 AAGATGTT-GCATGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG 1230
Db        392 LysMetIleAlaCysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArg 411

Qy       1231 ---TCCCATCAGATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCT 1286
Db        412 PheProSerGluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---Ser 430

Qy       1287 TTG------AAGTTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGT 1337
Db        431 LeuPheGluSerSerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysVal 450

Qy       1338 AATGAGATAAATAAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCA 1394
Db        451 MetGluIleAsnArgGluValGluGluProProLysLysProSerAlaPheLysProAla 470

Qy       1395 TTGAAAGCAAAA---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACA 1442
Db        471 IleGluMetGlnAsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ 488

Qy       1443 AACATTGAAGCAGATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAAT 1498
Db        489 ThrLeuArgAlaAspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsn 508

Qy       1499 TCTTGGGATTCTAGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAA 1552
Db        509 SerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLy 528

Qy       1553 GGCTACACATCAAAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAAT 1606
Db        528 sAlaThrHisGlnLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnL 548

Qy       1607 AAAGTGGTCTTCTGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT--- 1660
Db        548 ysAspGlyLeuLeuLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeu 567

Qy       1661 AGAATTGAAGACATGCAAACTTTCAAAGCGAGCCTCC---GGGAAGCCATCTGCT---TC 1714
Db        568 GluLeuLysAspMetGlnThrPheLysAlaGlu-ProProGlyLysProSerAlaPheGl 587

Qy       1715 GAGCCTCCACTGAAATCAAAAGTCTGTCCCAAATAAGCCTTGG---------AATGAAAA 1765
```


Db 587 uProAlaThrGluMetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlu-- 606
Qy 1766 ATGAAAAACATGGAGGCAGATGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAA 1822
Db 607 ----GlnThrTrpArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyrG 625
Qy 1823 GAAAA---TTCTGGGATACTGGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGT 1875
Db 625 luGluAsnSerTrpAspThr-GluSerLeuCysGluThrValSerGlnLysAspValCys 644
Qy 1876 TTACCAAGG---CTGCGCTCAAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT--- 1929
Db 645 LeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGlySer 664
Qy 1930 CCCTGTTAAAGTGGTCTTCTGAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAA 1986
Db 665 ProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLys 684
Qy 1987 GCCTAGA---ATTGATGACATGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGC 2040
Db 685 Ala-LeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 704
Qy 2041 CTTCGAGCCTCCATTGAAA---TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGA 2094
Db 704 aPheGluProAlaIleGluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLysA 724
Qy 2095 ATGAACAAACATTGAGGCAGATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTA 2148
Db 724 snGluGlnThrLeuArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyr 743
Qy 2149 TGAGAAAGTTCTTGGGATTCTGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG-- 2203
Db 744 GluGluSerSerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCy 763
Qy 2204 -TTTACCCAGGCTACACATCAAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTC 2259
Db 763 sLeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGluSe 783
Qy 2260 TCT-GATAATGAGGTTTTCTGAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAA 2315
Db 783 rProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLy 803
Qy 2316 GCCT---TGAATTGATGGCATGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTG 2369
Db 803 sAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 823
Qy 2370 CTTCGAGCCTGCATTGAAAT---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAG 2423
Db 823 aPheGluProAlaIle-GluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLys 842
Qy 2424 AATGAACAACATTGAGACAGATCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGG 2477
Db 843 AsnGluGlnThr---LeuArgAlaAspGlnMetPheProSerGluSerLysGlnLysLys 861
Qy 2478 TTGAGAAAA---TTCTGGGATTCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGG 2531
Db 862 ValGluGluAsnSerTrpAspSerGlu-SerLeuArgGluThrValSerGlnLysAsp-V 881
Qy 2532 TGTGTACC---CAGGCTACACACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAG 2585
Db 881 alCysValProLysAlaThrHisGlnLysGluMetAspLysIleSerGlyLysLeuGluA 901
Qy 2586 ATTCACTAGCCTATCAAAAT---CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAA 2639
Db 901 spSerThrSerLeuSerLysIleLeuAspThrValHisSerCysGluArgAlaArgGluL 921
Qy 2640 CTTCAAAAAGACACTGTGAACA-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTT 2695
Db 921 euGlnLysAspHisCysGluGlnArgThrGlyLysMetGluGlnMetLysLysLysPheC 941
Qy 2696 GTGTACTGAAAAAGAAC---TGTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCA 2749
Db 941 ysValLeuLysLysLysLeuSerGluAlaLysGluIleLysSerGln-LeuGluAsnGln 960
Qy 2750 AAAGTTAAATGGGAACAAGAG---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAG 2803
Db 961 LysValLysTrpGluGlnGluLeuCysSerValArgLeuThrLeuAsnGlnGluGluGlu 980
Qy 2804 AGAAGAGAGAAATGCCGATATATTAAAGAAAAAA-----------TTAGGAAGAATTAG 2851
Db 981 LysArg-ArgAsnAlaAspIleLeuAsnGluLysIleArgGluGluLeuGlyArgIle-G 1000
Qy 2852 AAGAATCGAGAGCAGCATAGAAAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTC 2908
Db 1000 luGluGlnHisArg------LysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeu 1017
Qy 2909 AGATACAAG---ATAAGAATTGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---AC 2962
Db 1018 ArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln-ValSerHisTh 1037
Qy 2963 TCATGAAATGAAA---ATTATTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA-- 3017
Db 1037 rHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIleAlaMe 1057
Qy 3018 -TGCAAAACTGGAATAGCCACACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-T 3072
Db 1057 tLeuLysLeuGluIleAlaThrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrP 1077
Qy 3073 TTGAGGACTTAAGATTTTAAAGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGA 3128
Db 1077 heGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuL 1097
Qy 3129 AAAGGAATCATTACTAAAAG---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATA 3182
Db 1097 ysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIle 1116
Qy 3183 GCTGAGAAACAATGCTCATTCTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGA 3236
Db 1117 AlaGlu-AsnThrMetLeuThrSerLysLeuLysGluLysGlnAspLysGluIleLeuGl 1136
Qy 3237 GGCAGAAATTGAATA-CACCATCCTGACTGGCTTCGCTGTACAAGC-----------CA 3283
Db 1136 uAlaGluIleGluSerHisHisPro---------ArgLeuAlaSerAlaValGlnAspHi 1153
Qy 3284 TGATCAATTGTGACATCAAGAAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGAT 3337
Db 1153 sAspGlnIleValThrSerArgLysSerGlnGluProAlaPheHisIleAlaGlyAspAl 1173
Qy 3338 GCTGTTTGCAAAGAAAATGAATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCT 3394
Db 1173 aCysLeuGlnArgLysMetAsnValAspValSerSerThrIleTyr-AsnAsnGluValL 1193
Qy 3395 CCATCACCACTTTC---TGAGCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAAT 3448
Db 1193 euHisGlnProLeuSerGluAlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnT 1213
Qy 3449 TAGCAGGAGATGTCTAAGAGA---AATACATTGGTTTCAGAACATCACAAAGAGA---CA 3502
Db 1213 yrAlaGlyAsp-AlaLeuArgGluAsnThrLeuValSerGluHis-AlaGlnArgAspGl 1232
Qy 3503 ACGTGAAAACAGTGTCAAT---GAAGGAAGCGAACACATGTATCAAAAC------GACAA 3553
Db 1232 nArgGluThrGlnCysGlnMetLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAs 1252
Qy 3554 GATAATGGAACAAACACCTGAACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAG 3613
Db 1252 nValAsnLysHisThrGluGlnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnS 1272
Qy 3614 CAAAAATTGTGGCTTCACAGCAATT----AGTCATGCACATAGAAAGCTGACAACAAAAG 3669
Db 1272 erLysAsnMetTrpLeuGlnGlnGlnLeuValHisAlaHisLys-----------LysA 1288
Qy 3670 CAGATAACAAT--------------GATATTCATTTCTTG---AGAGGAAATGCAACAC 3711
Db 1288 laAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLysMetGlnHisH 1308

```
Qy       3712 ATCTCCTAAAAGAGAAAAAGAGGAGATATT---AATTACAATACCATTTAAAAACCGTAT 3768
Db       1308 isLeuLeuLysGluLysAsnGlu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArg 1327

Qy       3769 ATT---CAATATGAAAAAGAGAAA 3789
Db       1328 IleTyrGlnTyrGluLysGluLys 1335
```

RESULT 2
ABJ05537
ID ABJ05537 standard; protein; 1341 AA.
XX
AC ABJ05537;
XX
DT 14-NOV-2002 (first entry)
XX
DE Breast cancer-associated protein 2.
XX
KW Breast cancer; breast cancer-associated gene sequence; drug development;
KW pharmacogenetics; biosensor development.
XX
OS Unidentified.
XX
PN WO200259377-A2.
XX
PD 01-AUG-2002.
XX
PF 24-JAN-2002; 2002WO-US002242.
XX
PR 24-JAN-2001; 2001US-0263965P.
PR 02-FEB-2001; 2001US-0265928P.
PR 09-APR-2001; 2001US-00829472.
PR 09-APR-2001; 2001US-0282698P.
PR 04-MAY-2001; 2001US-0288590P.
PR 29-MAY-2001; 2001US-0294443P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Gish KC, Afar D;
XX
DR WPI; 2002-583738/62.
DR N-PSDB; ABT07694.
XX
PT Detecting a breast cancer-associated transcript in a patient's cell,
PT useful for diagnosing breast cancer, comprises contacting a biological
PT sample with a polynucleotide that selectively hybridizes with breast
PT cancer nucleic acids.
XX
PS Disclosure; Page 348-349; 414pp; English.
XX
CC The invention comprises a method of detecting a breast cancer-associated
CC transcript in a cell from a patient. The method of the invention involves
CC contacting a biological sample from the patient with a nucleotide that
CC hybridises to one of the 69 breast cancer-associated gene sequences shown
CC in the specification. The method of the invention is useful in the
CC diagnosis or prognosis of breast cancer, and for detecting genes that are
CC up or down-regulated in breast cancer cells. Genes identified by the
CC method of the invention can be used in diagnostic purposes and also as
CC targets for screening for therapeutic compounds that modulate breast
CC cancer (e.g. hormones or antibodies). Identification of genes that are
CC over or under expressed in breast cancer can additionally provide high-
CC resolution, high-sensitivity datasets which can be used in the areas of
CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
CC structure and biosensor development. Amino acid sequences ABJ05536 -
CC ABJ05604 represent the proteins encoded by the 69 breast cancer-
CC associated genes of the invention
XX
SQ Sequence 1341 AA;

Alignment Scores:

Pred. No.:	1.04e-131	Length:	1341
Score:	1234.00	Matches:	679
Percent Similarity:	59.18%	Conservative:	146
Best Local Similarity:	48.71%	Mismatches:	376
Query Match:	17.41%	Indels:	207
DB:	5	Gaps:	115

US-09-602-362E-22 (1-4115) x ABJ05537 (1-1341)

```
Qy         96 CAAAGGGAAGAAGACATCAACCTTAATATACAAGAG---CCCAGAAGAGACTGCTCTAAC 152
Db          3 LysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArg-ThrAlaLeu-- 21

Qy        153 TGGGCCTGGTCAATGGCCT--------------GAGGAAGTAGTAACATTTC---GGTA 194
Db         22 ----------HisTrpAlaCysValAsnGlyHisGluGluValValThr-PheLeuValA 38

Qy        195 GACAGAAGTGCCAGCTGACGTCCTTG---TGGCGAACACGGACACCTCTGATGAAGGC-T 250
Db         38 spArgLysCysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaL 58

Qy        251 TACAATGCCACAGGAGGCTTTGC---AAATATTTGATAGATTC-GGTGCCGATATAAATC 306
Db         58 euGlnCysHisGlnGluAlaCysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnL 78

Qy        307 TCGTG---ATGTGTATGCAACATGGCTTCCATTATGCGTTTATAGTGGATTTTGTCAGT- 362
Db         78 euValAspValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuServ 98

Qy        363 --GGTGGCAAACTGCTGTC-CATGGTGCAGTATCGAAGTGC---CAACAAGGCTGCCTCA 416
Db         98 alValAlaLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuT 118

Qy        417 CACCACTTTTACTACCAT---AACGAAAGAAGTGAGCAATTGTGGAATTTT---GCTGAT 470
Db        118 hrProLeuLeuLeu-SerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIle 137

Qy        471 AAAAATGCAAATGCGAATGCGTTA---ATAAGTTAAATGCACACCCTCATGCTGCTGTAT 527
Db        138 LysAsnAlaAsnAlaAsnAlaValAsnLys-TyrLysCysThrAlaLeuMetLeuAlaVa 157

Qy        528 G---TCTGGATCATCAGAGATAGTTGCATGCTTCTTAGCAAAA---TGTGACGTCTTT-- 579
Db        157 lCysHisGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAl 177

Qy        580 -GTGCAGATATAGTGGAGTAACTGCAGAACAT---ATGCTGTTACTGTGGATTTCTCACA 635
Db        177 aAlaAspIle-CysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisI 197

Qy        636 TTCATAACAAA---TTATGAATATATACGAAAATTATCAAAAATCA---TCAAATACCAA 689
Db        197 leHisGluGlnIleMet-GluTyrIleArgLysLeuSer-LysAsnHisGlnAsnThrAs 216

Qy        690 T---CAGAAGGAACTCTGCAGGAAACCTGATGAGGCTGCACCCT---GGCGGAAAGACAC 743
Db        216 nProGluGlyThrSerAlaGlyThrProAspGluAlaAla-ProLeuAlaGluArgThrP 236

Qy        744 CTGACACGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCACCCT---GGTGGA 797
Db        236 roAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAla-ProLeuValGlu 255

Qy        798 AAGACACCTGACAGGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCATCT--- 851
Db        256 ArgThrProAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeu 275

Qy        852 TGGTGGAGGAACATCTGACAAATTCAATGTTGGAGAAAGG---ACATCTGGAAAGTTCGA 908
Db        276 Val-GluGlyThrSerAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGl 295

Qy        909 ACGTCAGC---AGAAAAACACCTAGGAAATTACGATCC---TGCAAAAAAACATCTGAGA 962
Db        295 uGln-SerAlaGluGluThrProArgGluIleThrSerProAlaLysGluThrSerGluL 315

Qy        963 AATTTACG---GGCCAGCAAAGGAAGACCTAGAAGATCGCAGGGAGAA---AAAGAAGAC 1016
Db        315 ysPheThrTrpPro-AlaLysGlyArgProArgLysIleAlaTrpGluLysLysGluAsp 334
```

```
Db       1037 rHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIleAlaMe 1057

Qy       3018 -TGCAAAACTGGAATAGCCACACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-T 3072

Db       1057 tLeuLysLeuGluIleAlaThrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrP 1077

Qy       3073 TTGAGGACTTAAGATTTTAAAGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGA 3128

Db       1077 heGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuL 1097

Qy       3129 AAAGGAATCATTACTAAAAG---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATA 3182

Db       1097 ysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIle 1116

Qy       3183 GCTGAGAAACAATGCTCATTCTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGA 3236

Db       1117 AlaGlu-AsnThrMetLeuThrSerLysLeuLysGluLysGlnAspLysGluIleLeuGl 1136

Qy       3237 GGCAGAAATTGAATA-CACCATCCTGACTGGCTTCGCTGTACAAGC-----------CA 3283

Db       1136 uAlaGluIleGluSerHisHisPro---------ArgLeuAlaSerAlaValGlnAspHi 1153

Qy       3284 TGATCAATTGTGACATCAAGAAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGAT 3337

Db       1153 sAspGlnIleValThrSerArgLysSerGlnGluProAlaPheHisIleAlaGlyAspAl 1173

Qy       3338 GCTGTTTGCAAAGAAAATGAATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCT 3394

Db       1173 aCysLeuGlnArgLysMetAsnValAspValSerSerThrIleTyr-AsnAsnGluValL 1193

Qy       3395 CCATCACCACTTTC---TGAGCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAAT 3448

Db       1193 euHisGlnProLeuSerGluAlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnT 1213

Qy       3449 TAGCAGGAGATGTCTAAGAGA---AATACATTGGTTTCAGAACATCACAAAGAGA---CA 3502

Db       1213 yrAlaGlyAsp-AlaLeuArgGluAsnThrLeuValSerGluHis-AlaGlnArgAspGl 1232

Qy       3503 ACGTGAAAACAGTGTCAAT---GAAGGAAGCGAACACATGTATCAAAAC------GACAA 3553

Db       1232 nArgGluThrGlnCysGlnMetLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAs 1252

Qy       3554 GATAATGGAACAAACACCTGAACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAG 3613

Db       1252 nValAsnLysHisThrGluGlnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnS 1272

Qy       3614 CAAAAATTGTGGCTTCACAGCAATT----AGTCATGCACATAGAAAGCTGACAACAAAAG 3669

Db       1272 erLysAsnMetTrpLeuGlnGlnGlnLeuValHisAlaHisLys-----------LysA 1288

Qy       3670 CAGATAACAAT--------------GATATTCATTTCTTG---AGAGGAAATGCAACAC 3711

Db       1288 laAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLysMetGlnHisH 1308

Qy       3712 ATCTCCTAAAAGAGAAAAAGAGGAGATATT---AATTACAATACCATTTAAAAACCGTAT 3768

Db       1308 isLeuLeuLysGluLysAsnGlu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArg 1327

Qy       3769 ATT---CAATATGAAAAAGAGAAA 3789

Db       1328 IleTyrGlnTyrGluLysGluLys 1335
```

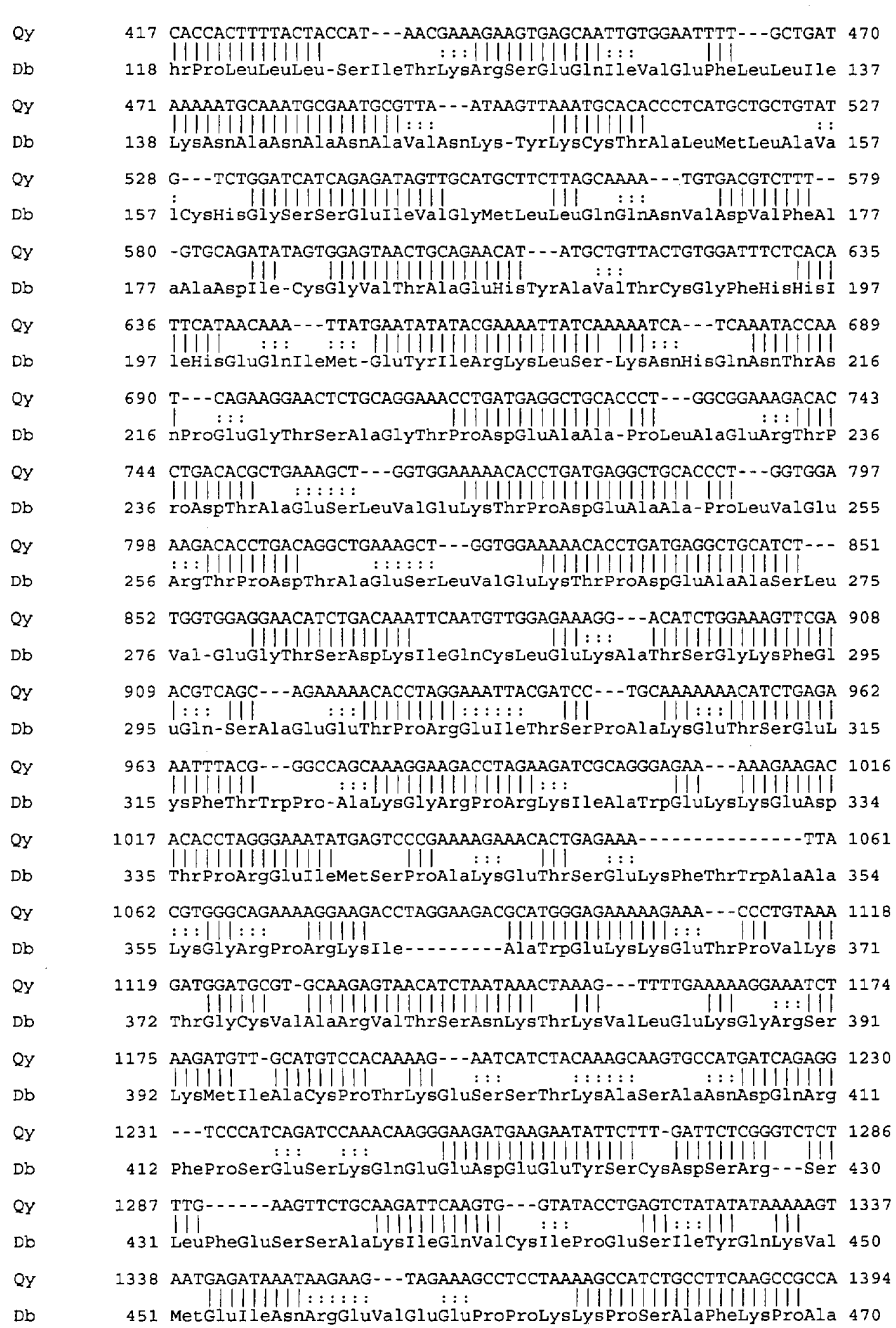
RESULT 3
ABJ37784
ID ABJ37784 standard; protein; 1341 AA.
XX
AC ABJ37784;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human tumour-related protein - SEQ ID No 565.
XX
KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX
OS Homo sapiens.
XX
PN WO200283956-A1.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, Mcneill PD, Durham M;
XX
DR WPI; 2003-103376/09.
XX
PT New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
PS Example 9; Page 342-346; 375pp; English.
XX
CC The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 1341 AA;

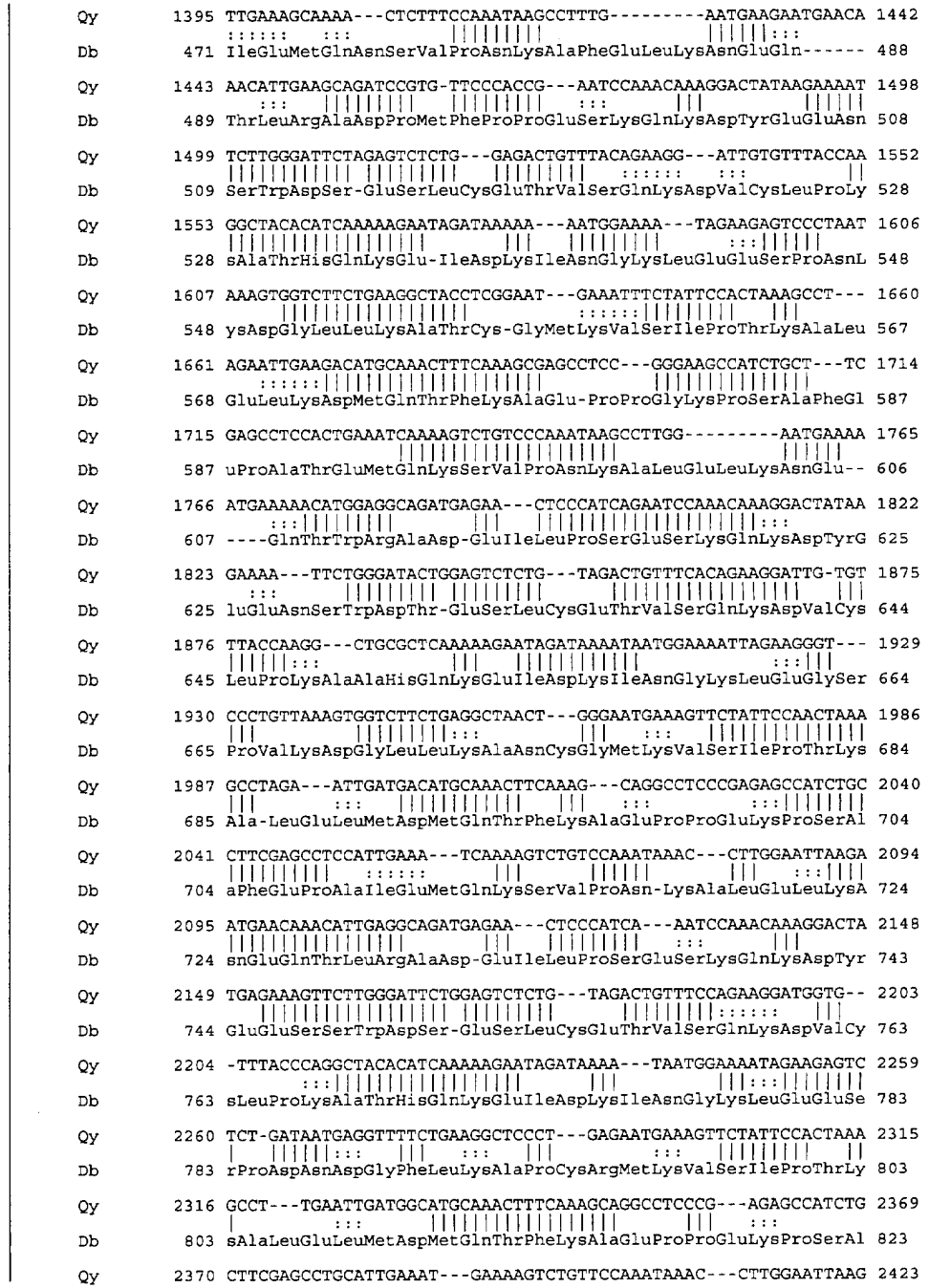
Alignment Scores:

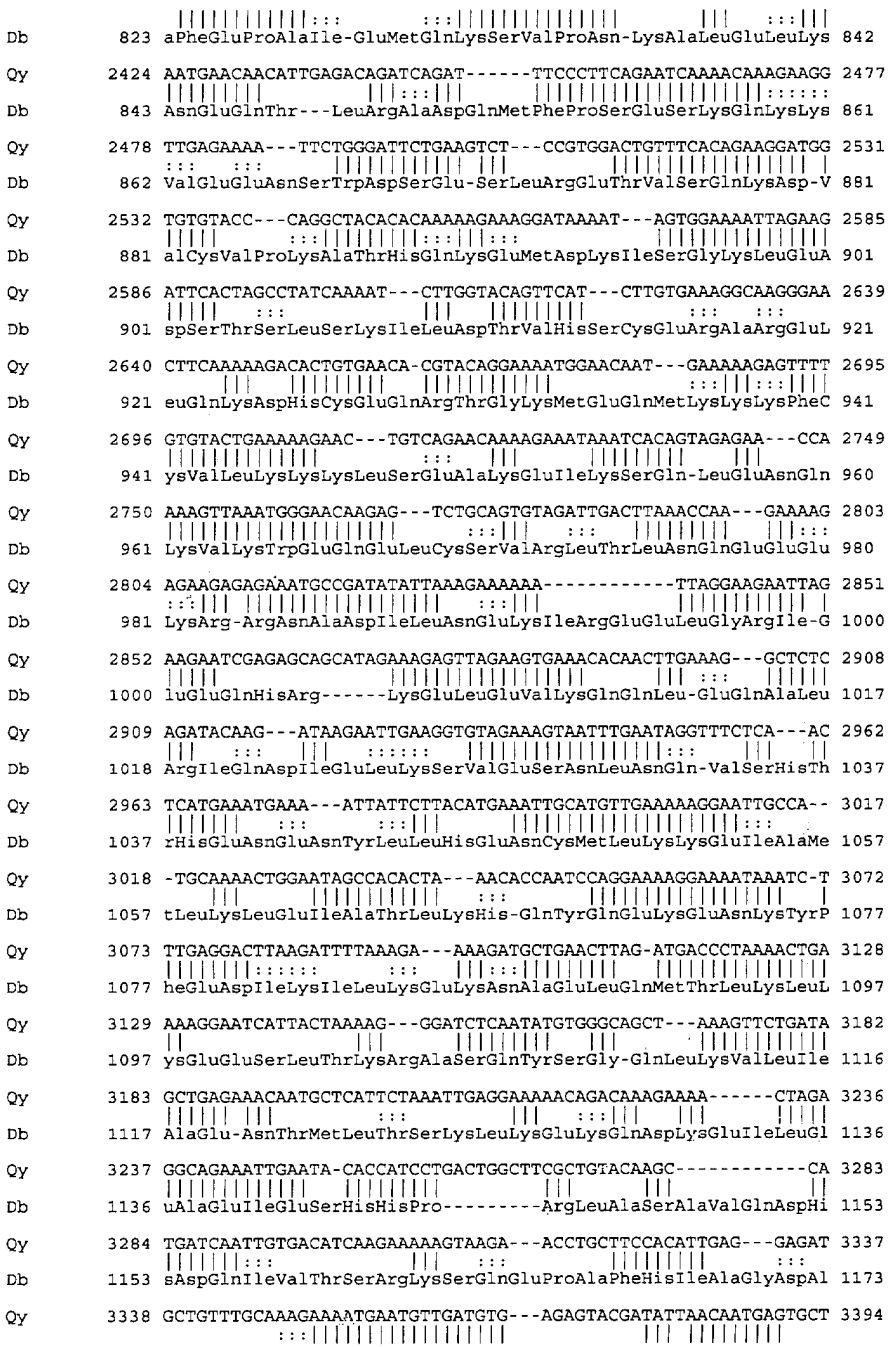
Pred. No.:	1.04e-131	Length:	1341
Score:	1234.00	Matches:	679
Percent Similarity:	59.18%	Conservative:	146
Best Local Similarity:	48.71%	Mismatches:	376
Query Match:	17.41%	Indels:	207
DB:	6	Gaps:	115

US-09-602-362E-22 (1-4115) x ABJ37784 (1-1341)

```
Qy         96 CAAAGGGAAGAAGACATCAACCTTAATATACAAGAG---CCCAGAAGAGACTGCTCTAAC 152

Db          3 LysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArg-ThrAlaLeu-- 21

Qy        153 TGGGCCTGGTCAATGGCCT--------------GAGGAAGTAGTAACATTTC---GGTA 194

Db         22 ----------HisTrpAlaCysValAsnGlyHisGluGluValValThr-PheLeuValA 38

Qy        195 GACAGAAGTGCCAGCTGACGTCCTTG---TGGCGAACACGGACACCTCTGATGAAGGC-T 250

Db         38 spArgLysCysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaL 58

Qy        251 TACAATGCCACAGGAGGCTTTGC---AAATATTTGATAGATTC-GGTGCCGATATAAATC 306

Db         58 euGlnCysHisGlnGluAlaCysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnL 78

Qy        307 TCGTG---ATGTGTATGCAACATGGCTTCCATTATGCGTTTATAGTGGATTTTGTCAGT- 362

Db         78 euValAspValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuServ 98

Qy        363 --GGTGGCAAACTGCTGTC-CATGGTGCAGTATCGAAGTGC---CAACAAGGCTGCCTCA 416

Db         98 alValAlaLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuT 118
```


Qy 417 CACCACTTTTACTACCAT---AACGAAAGAAGTGAGCAATTGTGGAATTTT---GCTGAT 470
Db 118 hrProLeuLeuLeu-SerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIle 137
Qy 471 AAAAATGCAAATGCGAATGCGTTA---ATAAGTTAAATGCACACCCTCATGCTGCTGTAT 527
Db 138 LysAsnAlaAsnAlaAsnAlaValAsnLys-TyrLysCysThrAlaLeuMetLeuAlaVa 157
Qy 528 G---TCTGGATCATCAGAGATAGTTGCATGCTTCTTAGCAAAA---TGTGACGTCTTT-- 579
Db 157 lCysHisGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAl 177
Qy 580 -GTGCAGATATAGTGGAGTAACTGCAGAACAT---ATGCTGTTACTGTGGATTTCTCACA 635
Db 177 aAlaAspIle-CysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisI 197
Qy 636 TTCATAACAAA---TTATGAATATATACGAAAATTATCAAAAATCA---TCAAATACCAA 689
Db 197 leHisGluGlnIleMet-GluTyrIleArgLysLeuSer-LysAsnHisGlnAsnThrAs 216
Qy 690 T---CAGAAGGAACTCTGCAGGAAACCTGATGAGGCTGCACCCT---GGCGGAAAGACAC 743
Db 216 nProGluGlyThrSerAlaGlyThrProAspGluAlaAla-ProLeuAlaGluArgThrP 236
Qy 744 CTGACACGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCACCCT---GGTGGA 797
Db 236 roAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAla-ProLeuValGlu 255
Qy 798 AAGACACCTGACAGGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCATCT--- 851
Db 256 ArgThrProAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeu 275
Qy 852 TGGTGGAGGAACATCTGACAAATTCAATGTTGGAGAAAGG---ACATCTGGAAAGTTCGA 908
Db 276 Val-GluGlyThrSerAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGl 295
Qy 909 ACGTCAGC---AGAAAAACACCTAGGAAATTACGATCC---TGCAAAAAAACATCTGAGA 962
Db 295 uGln-SerAlaGluGluThrProArgGluIleThrSerProAlaLysGluThrSerGluL 315
Qy 963 AATTTACG---GGCCAGCAAAGGAAGACCTAGAAGATCGCAGGGAGAA---AAAGAAGAC 1016
Db 315 ysPheThrTrpPro-AlaLysGlyArgProArgLysIleAlaTrpGluLysLysGluAsp 334
Qy 1017 ACACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAA--------------TTA 1061
Db 335 ThrProArgGluIleMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAla 354
Qy 1062 CGTGGGCAGAAAAGGAAGACCTAGGAAGACGCATGGGAGAAAAAGAAA---CCCTGTAAA 1118
Db 355 LysGlyArgProArgLysIle---------AlaTrpGluLysLysGluThrProValLys 371
Qy 1119 GATGGATGCGT-GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCT 1174
Db 372 ThrGlyCysValAlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSer 391
Qy 1175 AAGATGTT-GCATGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG 1230
Db 392 LysMetIleAlaCysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArg 411
Qy 1231 ---TCCCATCAGATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCT 1286
Db 412 PheProSerGluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---Ser 430
Qy 1287 TTG------AAGTTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGT 1337
Db 431 LeuPheGluSerSerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysVal 450
Qy 1338 AATGAGATAAATAAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCA 1394
Db 451 MetGluIleAsnArgGluValGluGluProProLysLysProSerAlaPheLysProAla 470


Qy 1395 TTGAAAGCAAAA---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACA 1442
Db 471 IleGluMetGlnAsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ 488
Qy 1443 AACATTGAAGCAGATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAAT 1498
Db 489 ThrLeuArgAlaAspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsn 508
Qy 1499 TCTTGGGATTCTAGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAA 1552
Db 509 SerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLy 528
Qy 1553 GGCTACACATCAAAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAAT 1606
Db 528 sAlaThrHisGlnLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnL 548
Qy 1607 AAAGTGGTCTTCTGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT--- 1660
Db 548 ysAspGlyLeuLeuLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeu 567
Qy 1661 AGAATTGAAGACATGCAAACTTTCAAAGCGAGCCTCC---GGGAAGCCATCTGCT---TC 1714
Db 568 GluLeuLysAspMetGlnThrPheLysAlaGlu-ProProGlyLysProSerAlaPheGl 587
Qy 1715 GAGCCTCCACTGAAATCAAAAGTCTGTCCCAAATAAGCCTTGG---------AATGAAAA 1765
Db 587 uProAlaThrGluMetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlu-- 606
Qy 1766 ATGAAAAACATGGAGGCAGATGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAA 1822
Db 607 ----GlnThrTrpArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyrG 625
Qy 1823 GAAAA---TTCTGGGATACTGGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGT 1875
Db 625 luGluAsnSerTrpAspThr-GluSerLeuCysGluThrValSerGlnLysAspValCys 644
Qy 1876 TTACCAAGG---CTGCGCTCAAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT--- 1929
Db 645 LeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGlySer 664
Qy 1930 CCCTGTTAAAGTGGTCTTCTGAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAA 1986
Db 665 ProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLys 684
Qy 1987 GCCTAGA---ATTGATGACATGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGC 2040
Db 685 Ala-LeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 704
Qy 2041 CTTCGAGCCTCCATTGAAA---TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGA 2094
Db 704 aPheGluProAlaIleGluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLysA 724
Qy 2095 ATGAACAAACATTGAGGCAGATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTA 2148
Db 724 snGluGlnThrLeuArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyr 743
Qy 2149 TGAGAAAGTTCTTGGGATTCTGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG-- 2203
Db 744 GluGluSerSerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCy 763
Qy 2204 -TTTACCCAGGCTACACATCAAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTC 2259
Db 763 sLeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGluSe 783
Qy 2260 TCT-GATAATGAGGTTTTCTGAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAA 2315
Db 783 rProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLy 803
Qy 2316 GCCT---TGAATTGATGGCATGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTG 2369
Db 803 sAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 823
Qy 2370 CTTCGAGCCTGCATTGAAAT---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAG 2423



RESULT 4
ABR47548
ID ABR47548 standard; protein; 1341 AA.
XX
AC ABR47548;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated protein sequence SEQ ID NO:334.
XX
KW Human; breast cancer; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003004989-A2.
XX
PD 16-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019669.
XX
PR 21-JUN-2001; 2001US-0299887P.
PR 27-JUN-2001; 2001US-0301572P.
PR 18-JUL-2001; 2001US-0306501P.
PR 25-SEP-2001; 2001US-0325002P.
PR 05-MAR-2002; 2002US-0362585P.
PR 14-MAY-2002; 2002US-0380391P.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
DR WPI; 2003-210381/20.
DR N-PSDB; ACC50246.
XX
PT Breast cancer diagnosis or treatment by comparing the level of expression
PT of a marker in a patient sample with that in the control non-breast
PT cancer sample.

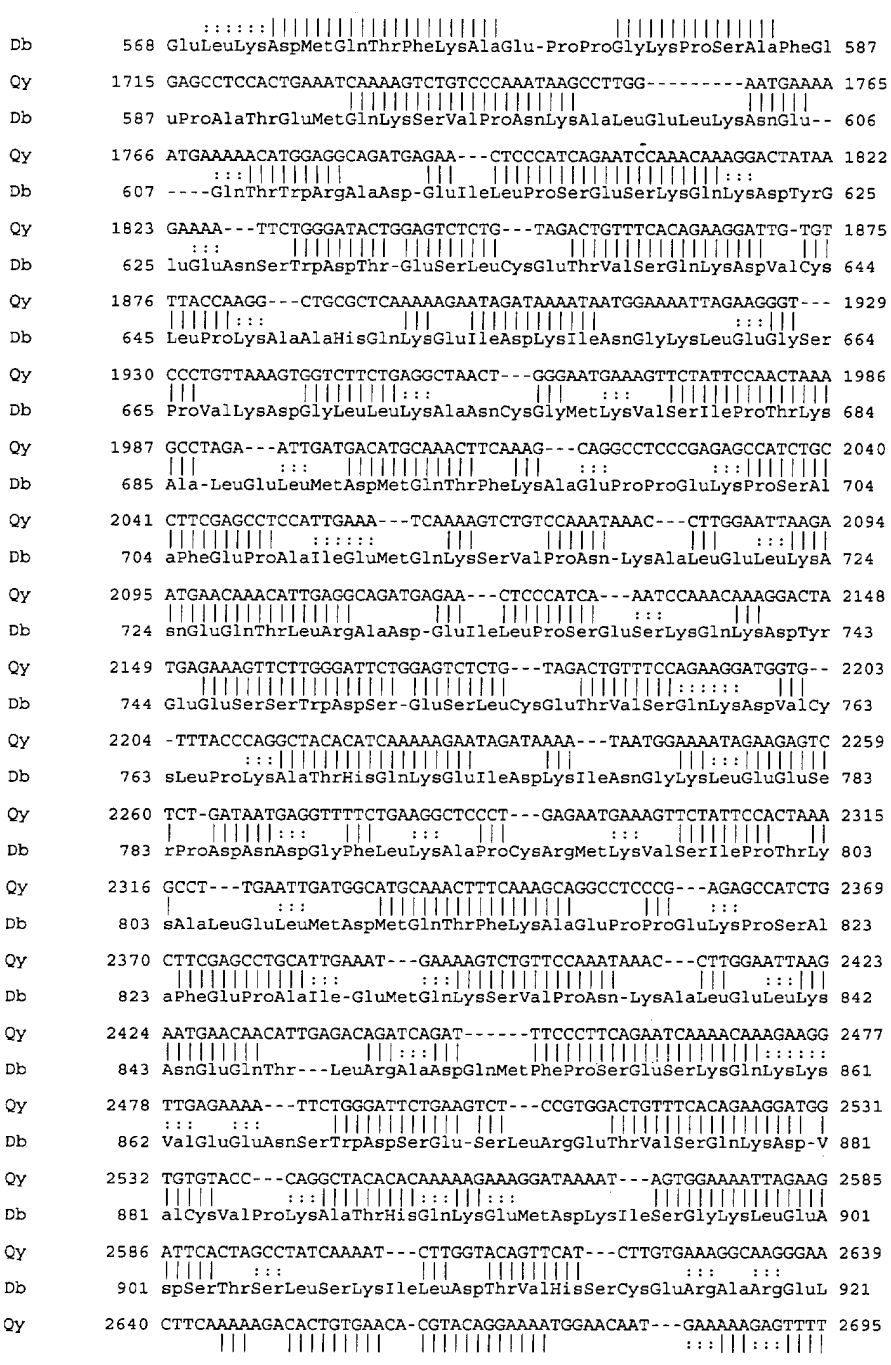
XX
PS Claim 1; SEQ ID NO 334; 128pp; English.
XX
CC The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1341 AA;

Alignment Scores:

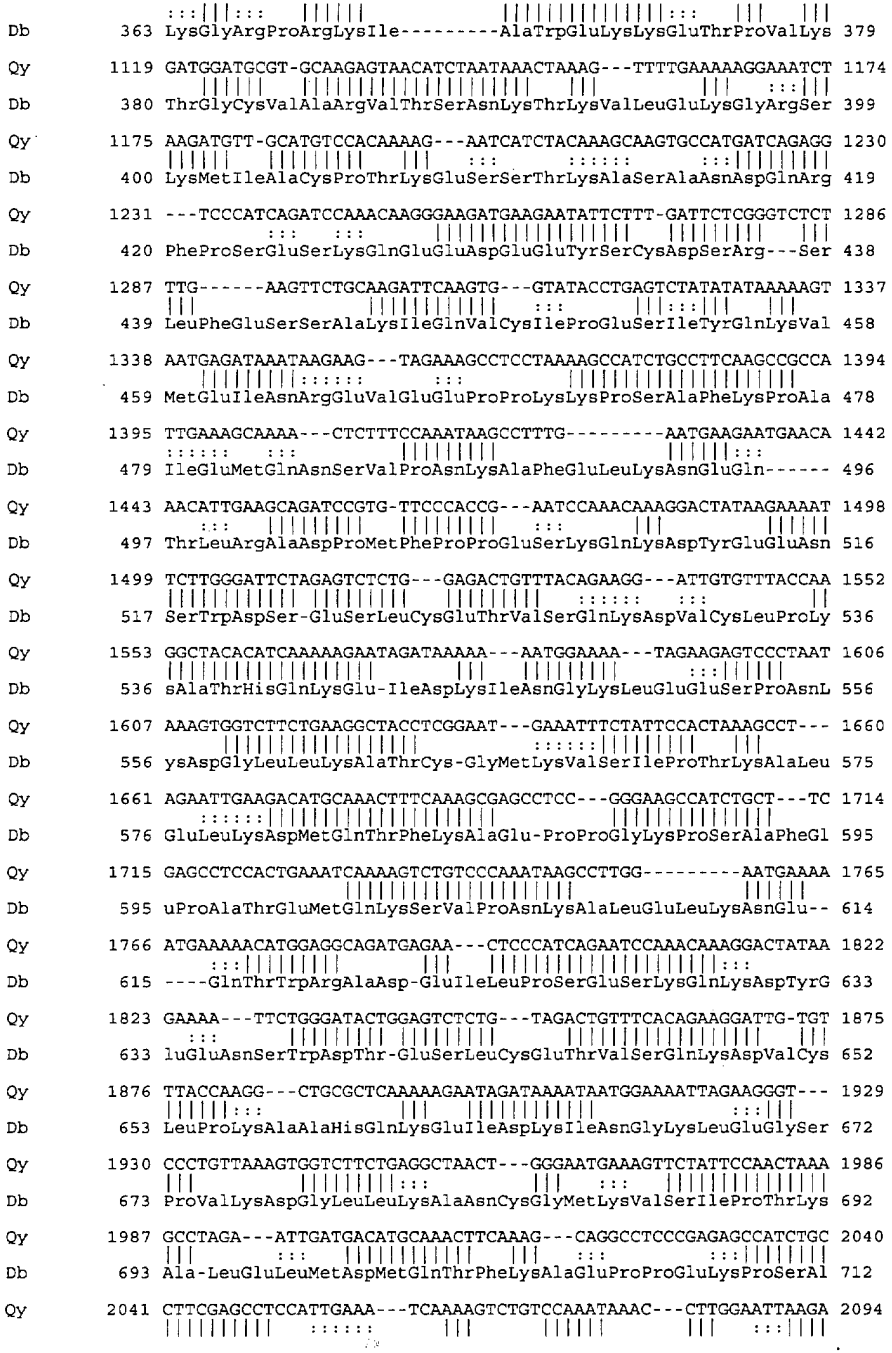
Pred. No.:	1.04e-131	Length:	1341
Score:	1234.00	Matches:	679
Percent Similarity:	59.18%	Conservative:	146
Best Local Similarity:	48.71%	Mismatches:	376
Query Match:	17.41%	Indels:	207
DB:	6	Gaps:	115

US-09-602-362E-22 (1-4115) x ABR47548 (1-1341)

```
Qy         96 CAAAGGGAAGAAGACATCAACCTTAATATACAAGAG---CCCAGAAGAGACTGCTCTAAC 152
Db          3 LysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArg-ThrAlaLeu-- 21

Qy        153 TGGGCCTGGTCAATGGCCT--------------GAGGAAGTAGTAACATTTC---GGTA 194
Db         22 ----------HisTrpAlaCysValAsnGlyHisGluGluValValThr-PheLeuValA 38

Qy        195 GACAGAAGTGCCAGCTGACGTCCTTG---TGGCGAACACGGACACCTCTGATGAAGGC-T 250
Db         38 spArgLysCysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaL 58

Qy        251 TACAATGCCACAGGAGGCTTTGC---AAATATTTGATAGATTC-GGTGCCGATATAAATC 306
Db         58 euGlnCysHisGlnGluAlaCysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnL 78

Qy        307 TCGTG---ATGTGTATGCAACATGGCTTCCATTATGCGTTTATAGTGGATTTTGTCAGT- 362
Db         78 euValAspValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerV 98

Qy        363 --GGTGGCAAACTGCTGTC-CATGGTGCAGTATCGAAGTGC---CAACAAGGCTGCCTCA 416
Db         98 alValAlaLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuT 118

Qy        417 CACCACTTTTACTACCAT---AACGAAAGAAGTGAGCAATTGTGGAATTTT---GCTGAT 470
Db        118 hrProLeuLeuLeu-SerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIle 137

Qy        471 AAAAATGCAAATGCGAATGCGTTA---ATAAGTTAAATGCACACCCTCATGCTGCTGTAT 527
Db        138 LysAsnAlaAsnAlaAsnAlaValAsnLys-TyrLysCysThrAlaLeuMetLeuAlaVa 157

Qy        528 G---TCTGGATCATCAGAGATAGTTGCATGCTTCTTAGCAAAA---TGTGACGTCTTT-- 579
Db        157 lCysHisGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAl 177

Qy        580 -GTGCAGATATAGTGGAGTAACTGCAGAACAT---ATGCTGTTACTGTGGATTTCTCACA 635
Db        177 aAlaAspIle-CysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisI 197

Qy        636 TTCATAACAAA---TTATGAATATATACGAAAATTATCAAAAATCA---TCAAATACCAA 689
Db        197 leHisGluGlnIleMet-GluTyrIleArgLysLeuSer-LysAsnHisGlnAsnThrAs 216

Qy        690 T---CAGAAGGAACTCTGCAGGAAACCTGATGAGGCTGCACCCT---GGCGGAAAGACAC 743
Db        216 nProGluGlyThrSerAlaGlyThrProAspGluAlaAla-ProLeuAlaGluArgThrP 236

Qy        744 CTGACACGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCACCCT---GGTGGA 797
Db        236 roAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAla-ProLeuValGlu 255

Qy        798 AAGACACCTGACAGGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCATCT--- 851
Db        256 ArgThrProAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeu 275

Qy        852 TGGTGGAGGAACATCTGACAAATTCAATGTTGGAGAAAGG---ACATCTGGAAAGTTCGA 908
Db        276 Val-GluGlyThrSerAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGl 295

Qy        909 ACGTCAGC---AGAAAAACACCTAGGAAATTACGATCC---TGCAAAAAAACATCTGAGA 962
Db        295 uGln-SerAlaGluGluThrProArgGluIleThrSerProAlaLysGluThrSerGluL 315

Qy        963 AATTTACG---GGCCAGCAAAGGAAGACCTAGAAGATCGCAGGGAGAA---AAAGAAGAC 1016
Db        315 ysPheThrTrpPro-AlaLysGlyArgProArgLysIleAlaTrpGluLysLysGluAsp 334

Qy       1017 ACACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAA--------------TTA 1061
Db        335 ThrProArgGluIleMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAla 354

Qy       1062 CGTGGGCAGAAAAGGAAGACCTAGGAAGACGCATGGGAGAAAAAGAAA---CCCTGTAAA 1118
Db        355 LysGlyArgProArgLysIle---------AlaTrpGluLysLysGluThrProValLys 371

Qy       1119 GATGGATGCGT-GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCT 1174
Db        372 ThrGlyCysValAlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSer 391

Qy       1175 AAGATGTT-GCATGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG 1230
Db        392 LysMetIleAlaCysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArg 411

Qy       1231 ---TCCCATCAGATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCT 1286
Db        412 PheProSerGluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---Ser 430

Qy       1287 TTG------AAGTTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGT 1337
Db        431 LeuPheGluSerSerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysVal 450

Qy       1338 AATGAGATAAATAAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCA 1394
Db        451 MetGluIleAsnArgGluValGluGluProProLysLysProSerAlaPheLysProAla 470

Qy       1395 TTGAAAGCAAAA---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACA 1442
Db        471 IleGluMetGlnAsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ 488

Qy       1443 AACATTGAAGCAGATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAAT 1498
Db        489 ThrLeuArgAlaAspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsn 508

Qy       1499 TCTTGGGATTCTAGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAA 1552
Db        509 SerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLy 528

Qy       1553 GGCTACACATCAAAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAAT 1606
Db        528 sAlaThrHisGlnLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnL 548

Qy       1607 AAAGTGGTCTTCTGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT--- 1660
Db        548 ysAspGlyLeuLeuLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeu 567

Qy       1661 AGAATTGAAGACATGCAAACTTTCAAAGCGAGCCTCC---GGGAAGCCATCTGCT---TC 1714
```


Db 568 GluLeuLysAspMetGlnThrPheLysAlaGlu-ProProGlyLysProSerAlaPheGl 587
Qy 1715 GAGCCTCCACTGAAATCAAAAGTCTGTCCCAAATAAGCCTTGG---------AATGAAAA 1765
Db 587 uProAlaThrGluMetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlu-- 606
Qy 1766 ATGAAAAACATGGAGGCAGATGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAA 1822
Db 607 ----GlnThrTrpArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyrG 625
Qy 1823 GAAAA---TTCTGGGATACTGGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGT 1875
Db 625 luGluAsnSerTrpAspThr-GluSerLeuCysGluThrValSerGlnLysAspValCys 644
Qy 1876 TTACCAAGG---CTGCGCTCAAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT--- 1929
Db 645 LeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGlySer 664
Qy 1930 CCCTGTTAAAGTGGTCTTCTGAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAA 1986
Db 665 ProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLys 684
Qy 1987 GCCTAGA---ATTGATGACATGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGC 2040
Db 685 Ala-LeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 704
Qy 2041 CTTCGAGCCTCCATTGAAA---TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGA 2094
Db 704 aPheGluProAlaIleGluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLysA 724
Qy 2095 ATGAACAAACATTGAGGCAGATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTA 2148
Db 724 snGluGlnThrLeuArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyr 743
Qy 2149 TGAGAAAGTTCTTGGGATTCTGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG-- 2203
Db 744 GluGluSerSerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCy 763
Qy 2204 -TTTACCCAGGCTACACATCAAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTC 2259
Db 763 sLeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGluSe 783
Qy 2260 TCT-GATAATGAGGTTTTCTGAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAA 2315
Db 783 rProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLy 803
Qy 2316 GCCT---TGAATTGATGGCATGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTG 2369
Db 803 sAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 823
Qy 2370 CTTCGAGCCTGCATTGAAAT---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAG 2423
Db 823 aPheGluProAlaIle-GluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLys 842
Qy 2424 AATGAACAACATTGAGACAGATCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGG 2477
Db 843 AsnGluGlnThr---LeuArgAlaAspGlnMetPheProSerGluSerLysGlnLysLys 861
Qy 2478 TTGAGAAAA---TTCTGGGATTCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGG 2531
Db 862 ValGluGluAsnSerTrpAspSerGlu-SerLeuArgGluThrValSerGlnLysAsp-V 881
Qy 2532 TGTGTACC---CAGGCTACACACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAG 2585
Db 881 alCysValProLysAlaThrHisGlnLysGluMetAspLysIleSerGlyLysLeuGluA 901
Qy 2586 ATTCACTAGCCTATCAAAAT---CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAA 2639
Db 901 spSerThrSerLeuSerLysIleLeuAspThrValHisSerCysGluArgAlaArgGluL 921
Qy 2640 CTTCAAAAAGACACTGTGAACA-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTT 2695
Db 921 euGlnLysAspHisCysGluGlnArgThrGlyLysMetGluGlnMetLysLysLysPheC 941
Qy 2696 GTGTACTGAAAAAGAAC---TGTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCA 2749
Db 941 ysValLeuLysLysLysLeuSerGluAlaLysGluIleLysSerGln-LeuGluAsnGln 960
Qy 2750 AAAGTTAAATGGGAACAAGAG---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAG 2803
Db 961 LysValLysTrpGluGlnGluLeuCysSerValArgLeuThrLeuAsnGlnGluGluGlu 980
Qy 2804 AGAAGAGAGAAATGCCGATATATTAAAGAAAAAA-----------TTAGGAAGAATTAG 2851
Db 981 LysArg-ArgAsnAlaAspIleLeuAsnGluLysIleArgGluGluLeuGlyArgIle-G 1000
Qy 2852 AAGAATCGAGAGCAGCATAGAAAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTC 2908
Db 1000 luGluGlnHisArg------LysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeu 1017
Qy 2909 AGATACAAG---ATAAGAATTGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---AC 2962
Db 1018 ArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln-ValSerHisTh 1037
Qy 2963 TCATGAAATGAAA---ATTATTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA-- 3017
Db 1037 rHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIleAlaMe 1057
Qy 3018 -TGCAAAACTGGAATAGCCACACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-T 3072
Db 1057 tLeuLysLeuGluIleAlaThrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrP 1077
Qy 3073 TTGAGGACTTAAGATTTTAAAGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGA 3128
Db 1077 heGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuL 1097
Qy 3129 AAAGGAATCATTACTAAAAG---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATA 3182
Db 1097 ysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIle 1116
Qy 3183 GCTGAGAAACAATGCTCATTCTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGA 3236
Db 1117 AlaGlu-AsnThrMetLeuThrSerLysLeuLysGluLysGlnAspLysGluIleLeuGl 1136
Qy 3237 GGCAGAAATTGAATA-CACCATCCTGACTGGCTTCGCTGTACAAGC------------CA 3283
Db 1136 uAlaGluIleGluSerHisHisPro---------ArgLeuAlaSerAlaValGlnAspHi 1153
Qy 3284 TGATCAATTGTGACATCAAGAAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGAT 3337
Db 1153 sAspGlnIleValThrSerArgLysSerGlnGluProAlaPheHisIleAlaGlyAspAl 1173
Qy 3338 GCTGTTTGCAAAGAAAATGAATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCT 3394
Db 1173 aCysLeuGlnArgLysMetAsnValAspValSerSerThrIleTyr-AsnAsnGluValL 1193
Qy 3395 CCATCACCACTTTC---TGAGCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAAT 3448
Db 1193 euHisGlnProLeuSerGluAlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnT 1213
Qy 3449 TAGCAGGAGATGTCTAAGAGA---AATACATTGGTTTCAGAACATCACAAAGAGA---CA 3502
Db 1213 yrAlaGlyAsp-AlaLeuArgGluAsnThrLeuValSerGluHis-AlaGlnArgAspGl 1232
Qy 3503 ACGTGAAAACAGTGTCAAT---GAAGGAAGCGAACACATGTATCAAAAC------GACAA 3553
Db 1232 nArgGluThrGlnCysGlnMetLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAs 1252
Qy 3554 GATAATGGAACAAACACCTGAACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAG 3613
Db 1252 nValAsnLysHisThrGluGlnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnS 1272
Qy 3614 CAAAAATTGTGGCTTCACAGCAATT----AGTCATGCACATAGAAAGCTGACAACAAAAG 3669
Db 1272 erLysAsnMetTrpLeuGlnGlnGlnLeuValHisAlaHisLys----------LysA 1288

```
Qy       3670 CAGATAACAAT--------------GATATTCATTTCTTG---AGAGGAAATGCAACAC 3711
Db       1288 laAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLysMetGlnHisH 1308

Qy       3712 ATCTCCTAAAAGAGAAAAAGAGGAGATATT---AATTACAATACCATTTAAAAACCGTAT 3768
Db       1308 isLeuLeuLysGluLysAsnGlu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArg 1327

Qy       3769 ATT---CAATATGAAAAAGAGAAA 3789
Db       1328 IleTyrGlnTyrGluLysGluLys 1335

RESULT 5
ABJ37788
ID   ABJ37788 standard; protein; 1349 AA.
XX
AC   ABJ37788;
XX
DT   15-MAY-2003  (first entry)
XX
DE   Human tumour-related protein - SEQ ID No 573.
XX
KW   Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW   tumour; breast cancer; cancer; immune response stimulation.
XX
OS   Homo sapiens.
XX
PN   WO200283956-A1.
XX
PD   24-OCT-2002.
XX
PF   15-APR-2002; 2002WO-US012378.
XX
PR   13-APR-2001; 2001US-00834759.
PR   07-DEC-2001; 2001US-00007805.
PR   13-FEB-2002; 2002US-00076622.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Houghton RL,  Sleath PR,  Persing DH,  Jiang Y,  Dillon DC;
PI   Mitcham JL,  Xu J,  Harlocker SL,  Hepler WT,  Henderson RA,  Fanger GR;
PI   Vedvick TS,  Mcneill PD,  Durham M;
XX
DR   WPI; 2003-103376/09.
XX
PT   New polypeptide and polynucleotide useful for stimulating and/or
PT   expanding T cells specific for a tumor protein and treating breast
PT   cancer.
XX
PS   Example 12; Page 353-357; 375pp; English.
XX
CC   The invention comprises a method of stimulating and/or expanding T cells
CC   specific for a tumour protein. The invention further comprises human
CC   nucleic acids and proteins that are associated with tumours (e.g. breast
CC   cancer). The method and sequences of the invention are useful for
CC   stimulating and/or expanding T cells specific for a tumour protein,
CC   detecting the presence of cancer, stimulating an immune response in a
CC   patient and treating breast cancer. The present amino acid sequence
CC   represents a human tumour-related protein
XX
SQ   Sequence 1349 AA;

Alignment Scores:
Pred. No.:              1.92e-130      Length:          1349
Score:                  1223.00        Matches:          677
Percent Similarity:     59.07%         Conservative:     150
Best Local Similarity:  48.36%         Mismatches:       380
Query Match:            17.26%         Indels:           207
DB:                     6              Gaps:             115

US-09-602-362E-22 (1-4115) x ABJ37788 (1-1349)

Qy         96 CAAAGGGAAGAAGACATCAACCTTAATATACAAGAG---CCCAGAAGAGACTGCTCTAAC 152
Db         11 LysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArg-ThrAlaLeu-- 29

Qy        153 TGGGCCTGGTCAATGGCCT--------------GAGGAAGTAGTAACATTTC---GGTA 194
Db         30 ----------HisTrpAlaCysValAsnGlyHisGluGluValValThr-PheLeuValA 46

Qy        195 GACAGAAGTGCCAGCTGACGTCCTTG---TGGCGAACACGGACACCTCTGATGAAGGC-T 250
Db         46 spArgLysCysGlnProAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaL 66

Qy        251 TACAATGCCACAGGAGGCTTTGC---AAATATTTGATAGATTC-GGTGCCGATATAAATC 306
Db         66 euGlnCysHisGlnGluAlaCysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnL 86

Qy        307 TCGTG---ATGTGTATGCAACATGGCTTCCATTATGCGTTTATAGTGGATTTTGTCAGT- 362
Db         86 euValAspValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerV 106

Qy        363 --GGTGGCAAACTGCTGTC-CATGGTGCAGTATCGAAGTGC---CAACAAGGCTGCCTCA 416
Db        106 alValAlaLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuT 126

Qy        417 CACCACTTTTACTACCAT---AACGAAAGAAGTGAGCAATTGTGGAATTTT---GCTGAT 470
Db        126 hrProLeuLeuLeu-SerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIle 145

Qy        471 AAAAATGCAAATGCGAATGCGTTA---ATAAGTTAAATGCACACCCTCATGCTGCTGTAT 527
Db        146 LysAsnAlaAsnAlaAsnAlaValAsnLys-TyrLysCysThrAlaLeuMetLeuAlaVa 165

Qy        528 G---TCTGGATCATCAGAGATAGTTGCATGCTTCTTAGCAAAA---TGTGACGTCTTT-- 579
Db        165 lCysHisGlyLeuSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAl 185

Qy        580 -GTGCAGATATAGTGGAGTAACTGCAGAACAT---ATGCTGTTACTGTGGATTTCTCACA 635
Db        185 aAlaAspIle-CysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisI 205

Qy        636 TTCATAACAAA---TTATGAATATATACGAAAATTATCAAAAATCA---TCAAATACCAA 689
Db        205 leHisGluGlnIleMet-GluTyrIleArgLysLeuSer-LysAsnHisGlnAsnThrAs 224

Qy        690 T---CAGAAGGAACTCTGCAGGAAACCTGATGAGGCTGCACCCT---GGCGGAAAGACAC 743
Db        224 nProGluGlyThrSerAlaGlyThrProAspGluAlaAla-ProLeuAlaGluArgThrP 244

Qy        744 CTGACACGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCACCCT---GGTGGA 797
Db        244 roAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAla-ProLeuValGlu 263

Qy        798 AAGACACCTGACAGGCTGAAAGCT---GGTGGAAAAACACCTGATGAGGCTGCATCT--- 851
Db        264 ArgThrProAspThrAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeu 283

Qy        852 TGGTGGAGGAACATCTGACAAATTCAATGTTGGAGAAAGG---ACATCTGGAAAGTTCGA 908
Db        284 Val-GluGlyThrSerAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGl 303

Qy        909 ACGTCAGC---AGAAAAACACCTAGGAAATTACGATCC---TGCAAAAAAACATCTGAGA 962
Db        303 uGln-SerAlaGluGluThrProArgGluIleThrSerProAlaLysGluThrSerGluL 323

Qy        963 AATTTACG---GGCCAGCAAAGGAAGACCTAGAAGATCGCAGGGAGAA---AAAGAAGAC 1016
Db        323 ysPheThrTrpPro-AlaLysGlyArgProArgLysIleAlaTrpGluLysLysGluAsp 342

Qy       1017 ACACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAA--------------TTA 1061
Db        343 ThrProArgGluIleMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAla 362

Qy       1062 CGTGGGCAGAAAAGGAAGACCTAGGAAGACGCATGGGAGAAAAAGAAA---CCCTGTAAA 1118
```


Db 363 LysGlyArgProArgLysIle---------AlaTrpGluLysLysGluThrProValLys 379
Qy 1119 GATGGATGCGT-GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCT 1174
Db 380 ThrGlyCysValAlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSer 399
Qy 1175 AAGATGTT-GCATGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG 1230
Db 400 LysMetIleAlaCysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArg 419
Qy 1231 ---TCCCATCAGATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCT 1286
Db 420 PheProSerGluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---Ser 438
Qy 1287 TTG------AAGTTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGT 1337
Db 439 LeuPheGluSerSerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysVal 458
Qy 1338 AATGAGATAAATAAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCA 1394
Db 459 MetGluIleAsnArgGluValGluGluProProLysLysProSerAlaPheLysProAla 478
Qy 1395 TTGAAAGCAAAA---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACA 1442
Db 479 IleGluMetGlnAsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ 496
Qy 1443 AACATTGAAGCAGATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAAT 1498
Db 497 ThrLeuArgAlaAspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsn 516
Qy 1499 TCTTGGGATTCTAGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAA 1552
Db 517 SerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLy 536
Qy 1553 GGCTACACATCAAAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAAT 1606
Db 536 sAlaThrHisGlnLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnL 556
Qy 1607 AAAGTGGTCTTCTGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT--- 1660
Db 556 ysAspGlyLeuLeuLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeu 575
Qy 1661 AGAATTGAAGACATGCAAACTTTCAAAGCGAGCCTCC---GGGAAGCCATCTGCT---TC 1714
Db 576 GluLeuLysAspMetGlnThrPheLysAlaGlu-ProProGlyLysProSerAlaPheGl 595
Qy 1715 GAGCCTCCACTGAAATCAAAAGTCTGTCCCAAATAAGCCTTGG---------AATGAAAA 1765
Db 595 uProAlaThrGluMetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlu-- 614
Qy 1766 ATGAAAAACATGGAGGCAGATGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAA 1822
Db 615 ----GlnThrTrpArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyrG 633
Qy 1823 GAAAA---TTCTGGGATACTGGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGT 1875
Db 633 luGluAsnSerTrpAspThr-GluSerLeuCysGluThrValSerGlnLysAspValCys 652
Qy 1876 TTACCAAGG---CTGCGCTCAAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT--- 1929
Db 653 LeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGlySer 672
Qy 1930 CCCTGTTAAAGTGGTCTTCTGAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAA 1986
Db 673 ProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLys 692
Qy 1987 GCCTAGA---ATTGATGACATGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGC 2040
Db 693 Ala-LeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 712
Qy 2041 CTTCGAGCCTCCATTGAAA---TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGA 2094
Db 712 aPheGluProAlaIleGluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLysA 732
Qy 2095 ATGAACAAACATTGAGGCAGATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTA 2148
Db 732 snGluGlnThrLeuArgAlaAsp-GluIleLeuProSerGluSerLysGlnLysAspTyr 751
Qy 2149 TGAGAAAGTTCTTGGGATTCTGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG-- 2203
Db 752 GluGluSerSerTrpAspSer-GluSerLeuCysGluThrValSerGlnLysAspValCy 771
Qy 2204 -TTTACCCAGGCTACACATCAAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTC 2259
Db 771 sLeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGluSe 791
Qy 2260 TCT-GATAATGAGGTTTTCTGAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAA 2315
Db 791 rProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLy 811
Qy 2316 GCCT---TGAATTGATGGCATGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTG 2369
Db 811 sAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSerAl 831
Qy 2370 CTTCGAGCCTGCATTGAAAT---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAG 2423
Db 831 aPheGluProAlaIle-GluMetGlnLysSerValProAsn-LysAlaLeuGluLeuLys 850
Qy 2424 AATGAACAACATTGAGACAGATCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGG 2477
Db 851 AsnGluGlnThr---LeuArgAlaAspGlnMetPheProSerGluSerLysGlnLysLys 869
Qy 2478 TTGAGAAAA---TTCTGGGATTCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGG 2531
Db 870 ValGluGluAsnSerTrpAspSerGlu-SerLeuArgGluThrValSerGlnLysAsp-V 889
Qy 2532 TGTGTACC---CAGGCTACACACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAG 2585
Db 889 alCysValProLysAlaThrHisGlnLysGluMetAspLysIleSerGlyLysLeuGluA 909
Qy 2586 ATTCACTAGCCTATCAAAAT---CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAA 2639
Db 909 spSerThrSerLeuSerLysIleLeuAspThrValHisSerCysGluArgAlaArgGluL 929
Qy 2640 CTTCAAAAAGACACTGTGAACA-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTT 2695
Db 929 euGlnLysAspHisCysGluGlnArgThrGlyLysMetGluGlnMetLysLysLysPheC 949
Qy 2696 GTGTACTGAAAAAGAAC---TGTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCA 2749
Db 949 ysValLeuLysLysLysLeuSerGluAlaLysGluIleLysSerGln-LeuGluAsnGln 968
Qy 2750 AAAGTTAAATGGGAACAAGAG---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAG 2803
Db 969 LysValLysTrpGluGlnGluLeuCysSerValArgLeuThrLeuAsnGlnGluGluGlu 988
Qy 2804 AGAAGAGAGAAATGCCGATATATTAAAGAAAAAA------------TTAGGAAGAATTAG 2851
Db 989 LysArg-ArgAsnAlaAspIleLeuAsnGluLysIleArgGluGluLeuGlyArgIle-G 1008
Qy 2852 AAGAATCGAGAGCAGCATAGAAAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTC 2908
Db 1008 luGluGlnHisArg------LysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeu 1025
Qy 2909 AGATACAAG---ATAAGAATTGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---AC 2962
Db 1026 ArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln-ValSerHisTh 1045
Qy 2963 TCATGAAATGAAA---ATTATTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA-- 3017
Db 1045 rHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIleAlaMe 1065
Qy 3018 -TGCAAAACTGGAATAGCCACACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-T 3072
Db 1065 tLeuLysLeuGluIleAlaThrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrP 1085

```
Qy       3073 TTGAGGACTTAAGATTTTAAAGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGA 3128
Db       1085 heGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuL 1105

Qy       3129 AAAGGAATCATTACTAAAAG---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATA 3182
Db       1105 ysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIle 1124

Qy       3183 GCTGAGAAACAATGCTCATTCTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGA 3236
Db       1125 AlaGlu-AsnThrMetLeuThrSerLysLeuLysGluLysGlnAspLysGluIleLeuGl 1144

Qy       3237 GGCAGAAATTGAATA-CACCATCCTGACTGGCTTCGCTGTACAAGC------------CA 3283
Db       1144 uAlaGluIleGluSerHisHisPro---------ArgLeuAlaSerAlaValGlnAspHi 1161

Qy       3284 TGATCAATTGTGACATCAAGAAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGAT 3337
Db       1161 sAspGlnIleValThrSerArgLysSerGlnGluProAlaPheHisIleAlaGlyAspAl 1181

Qy       3338 GCTGTTTGCAAAGAAAATGAATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCT 3394
Db       1181 aCysLeuGlnArgLysMetAsnValAspValSerSerThrIleTyr-AsnAsnGluValL 1201

Qy       3395 CCATCACCACTTTC---TGAGCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAAT 3448
Db       1201 euHisGlnProLeuSerGluAlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnT 1221

Qy       3449 TAGCAGGAGATGTCTAAGAGA---AATACATTGGTTTCAGAACATCACAAAGAGA---CA 3502
Db       1221 yrAlaGlyAsp-AlaLeuArgGluAsnThrLeuValSerGluHis-AlaGlnArgAspGl 1240

Qy       3503 ACGTGAAAACAGTGTCAAT---GAAGGAAGCGAACACATGTATCAAAAC------GACAA 3553
Db       1240 nArgGluThrGlnCysGlnMetLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAs 1260

Qy       3554 GATAATGGAACAAACACCTGAACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAG 3613
Db       1260 nValAsnLysHisThrGluGlnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnS 1280

Qy       3614 CAAAAATTGTGGCTTCACAGCAATT----AGTCATGCACATAGAAAGCTGACAACAAAAG 3669
Db       1280 erLysAsnMetTrpLeuGlnGlnGlnLeuValHisAlaHisLys------------LysA 1296

Qy       3670 CAGATAACAAT--------------GATATTCATTTCTTG---AGAGGAAATGCAACAC 3711
Db       1296 laAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLysMetGlnHisH 1316

Qy       3712 ATCTCCTAAAAGAGAAAAAGAGGAGATATT---AATTACAATACCATTTAAAAACCGTAT 3768
Db       1316 isLeuLeuLysGluLysAsnGlu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArg 1335

Qy       3769 ATT---CAATATGAAAAAGAGAAAGAGAAACAGAAACTCATG 3807
Db       1336 IleTyrGlnTyrGluLysGluLysAlaGluThrGluValIle 1349
```

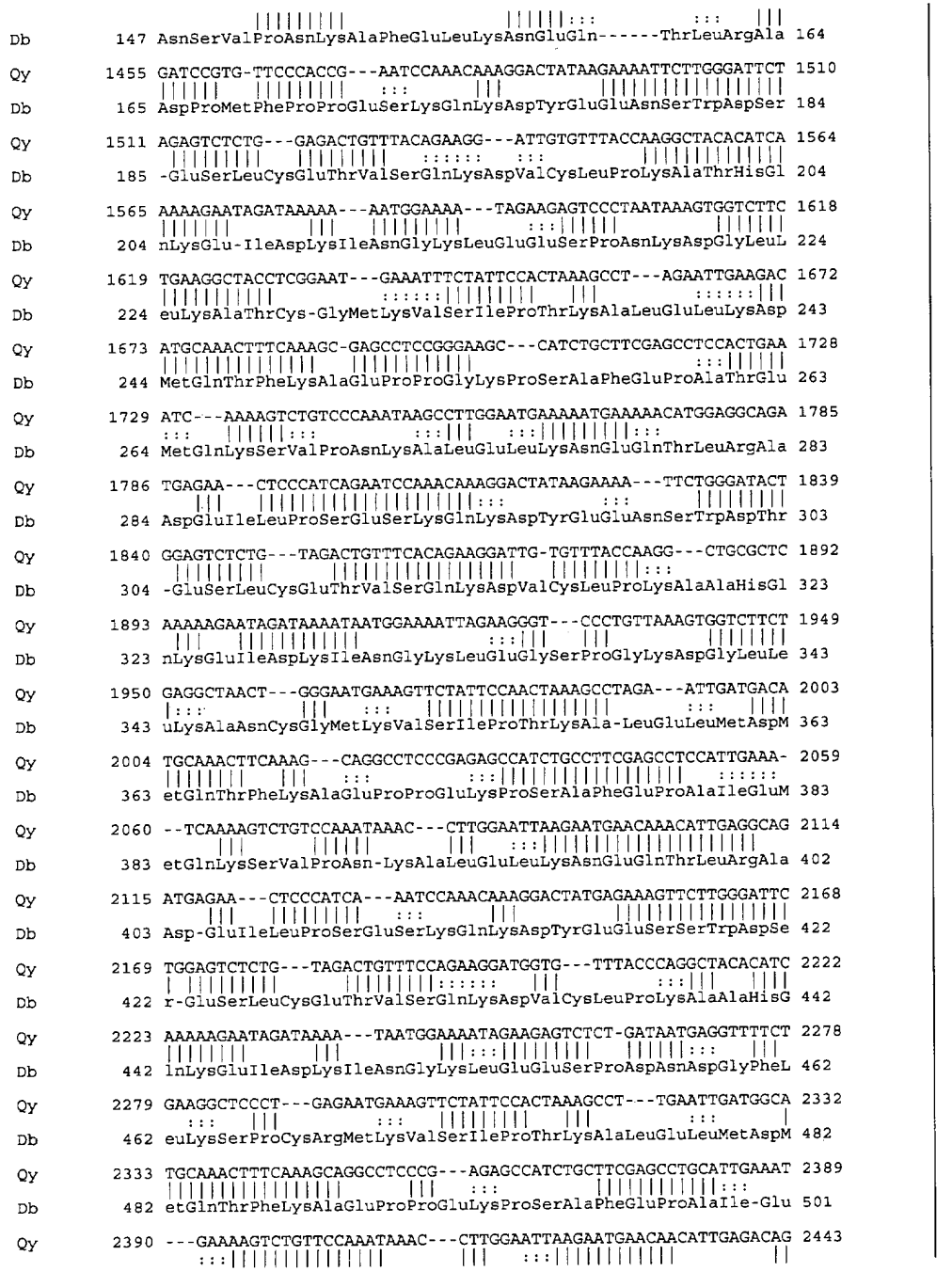
RESULT 6
ABJ37783
ID ABJ37783 standard; protein; 1013 AA.
XX
AC ABJ37783;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human tumour-related protein - SEQ ID No 553.
XX
KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX
OS Homo sapiens.
XX
PN WO200283956-A1.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, Mcneill PD, Durham M;
XX
DR WPI; 2003-103376/09.
XX
PT New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
PS Example 8; Page 335-338; 375pp; English.
XX
CC The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 1013 AA;

Alignment Scores:

Pred. No.:	2.31e-96	Length:	1013
Score:	926.50	Matches:	504
Percent Similarity:	59.39%	Conservative:	116
Best Local Similarity:	48.28%	Mismatches:	293
Query Match:	13.08%	Indels:	144
DB:	6	Gaps:	83

US-09-602-362E-22 (1-4115) x ABJ37783 (1-1013)

```
Qy       1018 CACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAATTACGTG---GGCAGAAAA 1074
Db          8 HisValGlySerMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLys 27

Qy       1075 GGAAGACCTAGGAAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT- 1129
Db         28 GlyArgProArgLysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysVal 47

Qy       1130 GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCA 1185
Db         48 AlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAla 67

Qy       1186 TGTCCACAAAAG---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAG 1239
Db         68 CysProThrLysGluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGlu 87

Qy       1240 ATCCAAACAAGGGAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAG 1292
Db         88 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSer 106

Qy       1293 TTCTGCAAGATTCAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAAT 1349
Db        107 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 126

Qy       1350 AAGAAG---TAGAAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA 1406
Db        127 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 146

Qy       1407 ---CTCTTTCCAAATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCA 1454
```


Db 147 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAla 164
Qy 1455 GATCCGTG-TTCCCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCT 1510
Db 165 AspProMetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer 184
Qy 1511 AGAGTCTCTG---GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCA 1564
Db 185 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGl 204
Qy 1565 AAAAGAATAGATAAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTC 1618
Db 204 nLysGlu-IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuL 224
Qy 1619 TGAAGGCTACCTCGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGAC 1672
Db 224 euLysAlaThrCys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAsp 243
Qy 1673 ATGCAAACTTTCAAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAA 1728
Db 244 MetGlnThrPheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGlu 263
Qy 1729 ATC---AAAAGTCTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGA 1785
Db 264 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 283
Qy 1786 TGAGAA---CTCCCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACT 1839
Db 284 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspThr 303
Qy 1840 GGAGTCTCTG---TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTC 1892
Db 304 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 323
Qy 1893 AAAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCT 1949
Db 323 nLysGluIleAspLysIleAsnGlyLysLeuGluGlySerProGlyLysAspGlyLeuLe 343
Qy 1950 GAGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACA 2003
Db 343 uLysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspM 363
Qy 2004 TGCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA- 2059
Db 363 etGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluM 383
Qy 2060 --TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAG 2114
Db 383 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 402
Qy 2115 ATGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTC 2168
Db 403 Asp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSe 422
Qy 2169 TGGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATC 2222
Db 422 r-GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisG 442
Qy 2223 AAAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCT 2278
Db 442 lnLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheL 462
Qy 2279 GAAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCA 2332
Db 462 euLysSerProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspM 482
Qy 2333 TGCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT 2389
Db 482 etGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-Glu 501
Qy 2390 ---GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAG 2443
Db 502 MetGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuAr 520
Qy 2444 ATCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGA 2494
Db 520 gAlaAspGlnMetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAs 540
Qy 2495 TTCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACA 2548
Db 540 pSerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThr 559
Qy 2549 CACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT 2605
Db 560 HisGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLys 579
Qy 2606 ---CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAA 2659
Db 580 IleLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGlu 599
Qy 2660 CA-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC--- 2712
Db 600 GlnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeu 619
Qy 2713 TGTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGA 2769
Db 620 SerGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGl 639
Qy 2770 G---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATA 2823
Db 639 uLeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspI 659
Qy 2824 TATTAAAGAAAAAA-----------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAG 2871
Db 659 leLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------ 676
Qy 2872 AAAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAAT 2925
Db 677 LysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLe 696
Qy 2926 TGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTA 2979
Db 696 uLysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrL 716
Qy 2980 TTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCA 3036
Db 716 euLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaT 736
Qy 3037 CACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAA 3092
Db 736 hrLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeu 755
Qy 3093 AGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG 3148
Db 756 LysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLys 775
Qy 3149 ---GGATCTCAATATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATT 3202
Db 776 ArgAlaSerGlnTyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuT 795
Qy 3203 CTAAATTGAGGAAAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACC 3255
Db 795 hrSerLysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisH 815
Qy 3256 ATCCTGACTGGCTTCGCTGTACAAGC-----------CATGATCAATTGTGACATCAAG 3303
Db 815 isPro---------ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerA 832
Qy 3304 AAAAAGTAAGA---ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGA 3357
Db 832 rgLysSerGlnGluProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetA 852
Qy 3358 ATGTTGATGTG---AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGA 3411
Db 852 snValAspValSerSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGlu 871

```
Qy       3412 GCTCAAAGGAAATCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAG 3468
Db        872 AlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuAr 891

Qy       3469 A---AATACATTGGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT- 3521
Db        891 gGluAsnThrLeuValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnM 911

Qy       3522 --GAAGGAAGCGAACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTG 3573
Db        911 etLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluG 931

Qy       3574 AACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAG 3633
Db        931 lnGlnGluSerLeu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGln 950

Qy       3634 CAATT----AGTCATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------- 3680
Db        951 GlnGlnLeuValHisAlaHisLys-----------LysAlaAspAsnLysSerLysIle 966

Qy       3681 ------GATATTCATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAG 3731
Db        967 ThrIleAspIleHisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsn 986

Qy       3732 AGGAGATATT---AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGA 3785
Db        987 Glu-GluIlePheAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGl 1006

Qy       3786 GAAAGAGAAACAGAAACTCATG 3807
Db       1006 uLysAlaGluThrGluValIle 1013
```

RESULT 7
AAU33351
ID AAU33351 standard; protein; 1002 AA.
XX
AC AAU33351;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human breast cancer protein B726P from alternatively spliced cDNA.
XX
KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy; immunogen.
XX
OS Homo sapiens.
XX
PN WO200179286-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US012164.
XX
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX
DR WPI; 2001-611721/70.
DR N-PSDB; AAS47411.
XX
PT Breast Tumor Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer.
XX
PS Claim 3; Page 281-283; 297pp; English.
XX
CC The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
XX
SQ Sequence 1002 AA;

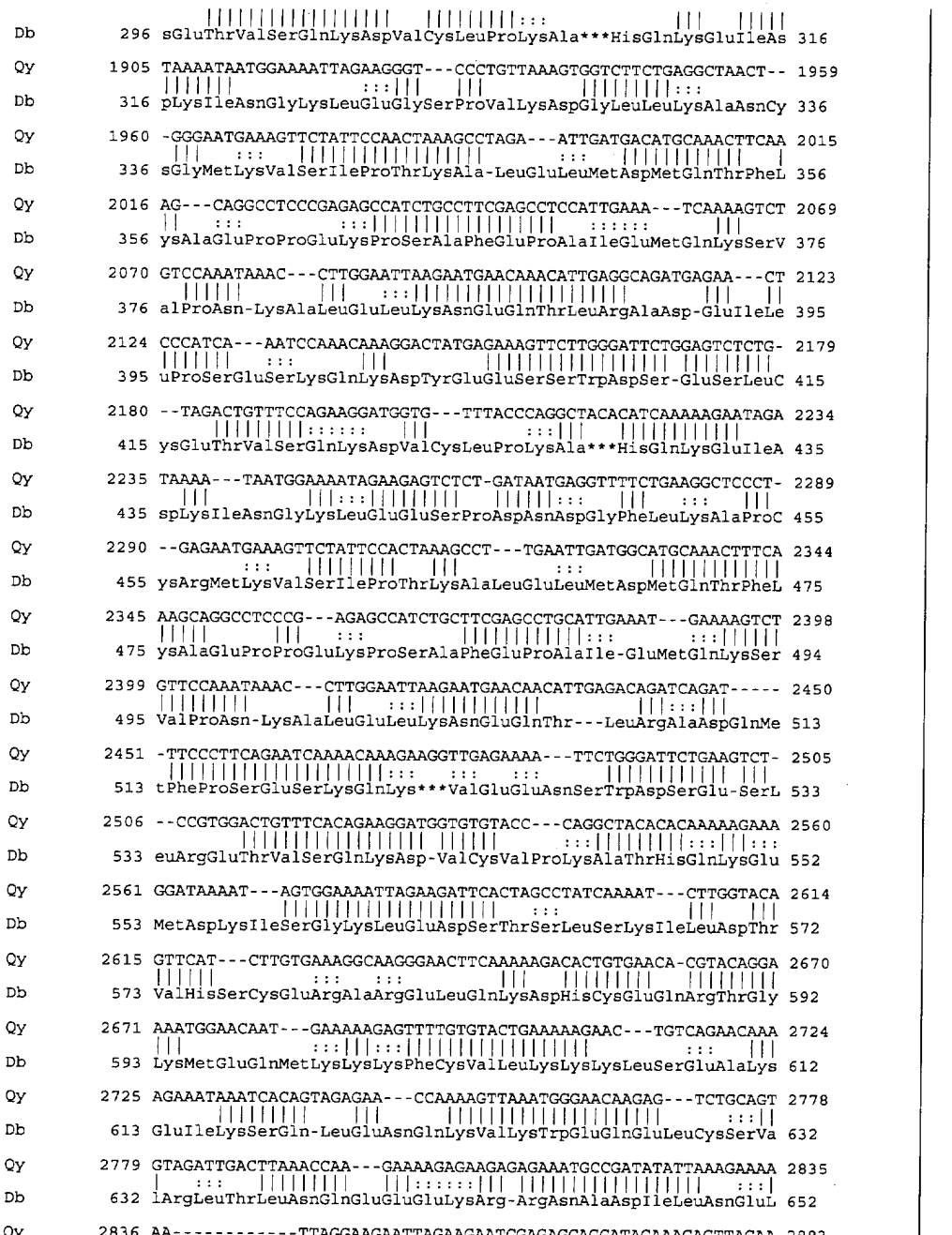
Alignment Scores:

Pred. No.:	2.85e-95	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	4	Gaps:	83

US-09-602-362E-22 (1-4115) x AAU33351 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
```

```
Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
              ||    |||||||||||||||||||||:::|||   :::      ||||||
Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
              |||||||||:::           ||||||||||||                  ||||||
Db        944 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
              |||||||||   |||          ||||||||||||||||||       :::|||
Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
               |||||||||         |||      |||   ||||||||||||||||||
Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996

RESULT 8
ABG78918
ID   ABG78918 standard; protein; 1002 AA.
XX
AC   ABG78918;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human breast tumour polypeptide #10.
XX
KW   Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   US2002085998-A1.
XX
PD   04-JUL-2002.
XX
PF   13-APR-2001; 2001US-00834759.
XX
PR   28-DEC-1998;   98US-00222575.
PR   02-APR-1999;   99US-00285480.
PR   23-JUN-1999;   99US-00339338.
PR   02-SEP-1999;   99US-00389681.
PR   03-NOV-1999;   99US-00433826.
PR   17-APR-2000; 2000US-00551621.
PR   08-JUN-2000; 2000US-00590751.
PR   22-JUN-2000; 2000US-00604287.
PR   20-JUL-2000; 2000US-00620405.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Dillon DC,  Mitcham JL,  Xu J,  Harlocker SL,  Hepler WT;
PI   Henderson RA;
XX
DR   WPI; 2002-635657/68.
DR   N-PSDB; ABS64012.
XX
PT   Novel breast cancer polynucleotides and polypeptides encoded by the
PT   polynucleotides, useful for detecting the presence of breast cancer in a
PT   patient, and in pharmaceutical compositions, for treating breast cancer.
XX
PS   Claim 2; Page 208-211; 247pp; English.
XX
CC   The invention relates to an isolated breast tumour polynucleotide and the
CC   polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC   detecting the presence of breast cancer in a patient, and in
CC   pharmaceutical compositions for treating breast cancer. The sequences are
CC   useful for stimulating an immune response in a patient and can therefore
CC   be used in production of vaccines. The sequences are also useful for
CC   detecting the presence of a cancer in a patient, by obtaining a
CC   biological sample from the patient, contacting the biological sample with
CC   a composition of the invention and detecting the amount of polynucleotide
CC   that hybridizes to the sample. This sequence represents a human breast
CC   tumour polypeptide of the invention
XX
SQ   Sequence 1002 AA;

Alignment Scores:
Pred. No.:                2.85e-95      Length:          1002
Score:                    917.00        Matches:         506
Percent Similarity:       59.77%        Conservative:    112
Best Local Similarity:    48.94%        Mismatches:      285
Query Match:              12.94%        Indels:          145
DB:                       5             Gaps:            83

US-09-602-362E-22 (1-4115) x ABG78918 (1-1002)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
              |||||||||   |||:::      |||      |||||||||   ||||||||||||||||
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
              |||   |||||||||||||||:::   |||   |||   ||||||   ||||||||||||
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
              |||||||||   |||          |||   :::|||||||||   |||||||||   |||
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
                  :::      ::::::      :::||||||||          :::   :::
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
              ||||||||||||||||||   |||||||||   ||||||                ||||||
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
              ||||||   :::      ||:::|||   |||      |||||||||::::::
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
              :::        ||||||||||||||||||||   ::::::   :::        |||
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
              ||||||               ||||||:::         :::   ||||||||   |||
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
              ||||||   :::      |||          |||||||||||||||||| |||||||||
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
               ||||||||   ::::::   :::          ||||||||||||||||||||
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
                |||   ||||||||      :::||||||          ||||||||||||||||||
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
                    ::::::|||||||||   |||        ::::::||||||||||||||||
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
              |||   ||||||||||||                  :::|||||:::   ||||||
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
              :::         :::|||   :::|||||||||:::                |||   |||
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
              |||||||||||||||||:::         :::      ||||||::: |||||||||
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
```



RESULT 9
ABJ37741
ID ABJ37741 standard; protein; 1002 AA.

XX
AC ABJ37741;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human tumour-related protein - SEQ ID No 475.
XX
KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX
OS Homo sapiens.
XX
PN WO200283956-A1.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, Mcneill PD, Durham M;
XX
DR WPI; 2003-103376/09.
XX
PT New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
PS Example 1; Page 298-300; 375pp; English.
XX
CC The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 1002 AA;

Alignment Scores:

Pred. No.:	2.85e-95	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	6	Gaps:	83

US-09-602-362E-22 (1-4115) x ABJ37741 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
             |||||||||  |||:::     |||      |||||||||  ||||||||||||||||
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
             |||  ||||||||||||||:::  |||  |||  ||||||  ||||||||||||
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
             |||||||||  |||         |||  :::|||||||||  |||||||||  |||
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
                :::     ::::::     :::|||||||||         :::   :::
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
             ||||||||||||||||||   |||||||||  ||||||                ||||||
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
             ||||||   :::     |||:::|||   |||        |||||||||::::::
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
                :::   ||||||||||||||||||||||   ::::::   :::        |||
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
             ||||||          ||||||:::        :::   |||||||||   |||
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
             ||||||   :::     |||          ||||||||||||||||||  |||||||||
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
              |||||||||   ::::::   :::         ||||||||||||||||||||
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
                |||   |||||||||      :::||||||          ||||||||||||||||
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
                ::::::||||||||   |||         ::::::||||||||||||||||
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
              |||   ||||||||||||                    :::||||||:::   ||||||
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
              :::      :::|||  :::|||||||||:::                |||   |||
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
             |||||||||||||||||:::        :::      ||||||::: |||||||||
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
              ||||||||||||||||||  |||||||||:::                |||   |||||
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
             |||||||             :::|||   |||        |||||||||:::
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
              |||   :::   ||||||||||||||||||       :::   ||||||||||||   |
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
              ||   :::          :::||||||||||||||||||   ::::::      |||
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
              ||||||       |||   :::||||||||||||||||||||||   |||   ||
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
              |||||||   :::      |||          ||||||||||||||||  |||||||||
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
                |||||||||::::::   |||         :::|||   ||||||||||||
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
```

```
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344

Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398

Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450

Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505

Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560

Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614

Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670

Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724

Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778

Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysServa 632

Qy       2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835

Db        632 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 652

Qy       2836 AA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883

Db        652 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 669

Qy       2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937

Db        670 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 689

Qy       2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991

Db        689 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 709

Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045

Db        709 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 728

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101

Db        729 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 748

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157

Db        749 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 768

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214

Db        769 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 788

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267

Db        788 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 805

Qy       3268 TTCGCTGTACAAGC-----------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314

Db        806 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 825

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368

Db        825 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 845

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423

Db        845 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 864

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477

Db        865 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 884

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531

Db        884 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 904

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585

Db        904 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 924

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641

Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686

Db        944 HisAlaHisLys------------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741

Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789

Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996
```

RESULT 10
ADE44427
ID ADE44427 standard; protein; 1002 AA.
XX
AC ADE44427;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human breast cancer protein #10.
XX
KW human; breast tumour; cancer; vaccine; T cell stimulator;
KW T cell expander.
XX
OS Homo sapiens.
XX
PN US2003104366-A1.
XX
PD 05-JUN-2003.
XX
PF 17-APR-2000; 2000US-00551621.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
XX
PA (JIAN/) JIANG Y.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (XUJJ/) XU J.

PA (HARL/) HARLOCKER S L.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX
DR WPI; 2004-020270/02.
DR N-PSDB; ADE44426.
XX
PT Novel isolated polypeptide comprising immunogenic portion of breast tumor
PT protein or its variant, useful for formulating vaccines for inhibiting
PT cancer development in a patient.
XX
PS Example 1; SEQ ID NO 475; 217pp; English.
XX
CC The invention relates to an isolated polypeptide comprising at least an
CC immunogenic portion of a breast tumour protein. The polynucleotide, its
CC polypeptide, its antibody, a pharmaceutical composition comprising the
CC fusion protein or the polynucleotide encoding it, a vaccine comprising
CC the fusion protein or the polynucleotide encoding it, an isolated T cell
CC population comprising T cells specific for a breast tumour protein, and a
CC method for removing tumour cells from a biological sample is useful for
CC inhibiting the development of a cancer in a patient. The polypeptide is
CC useful for stimulating and/or expanding T cells specific for a breast
CC tumour protein. Stimulating and/or expanding T cells specific for a
CC breast tumour protein is useful for inhibiting the development of a
CC cancer in a patient. The method additionally involves the step of cloning
CC at least one proliferated cell and then administering the cloned T cells
CC to the patient. The present sequence represents a breast cancer protein.
XX
SQ Sequence 1002 AA;

Alignment Scores:

Pred. No.:	2.85e-95	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	8	Gaps:	83

US-09-602-362E-22 (1-4115) x ADE44427 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
             |||||||||   |||:::     |||     |||||||||  ||||||||||||||||
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
             |||  ||||||||||||||:::   |||  |||   ||||||  ||||||||||||
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
             |||||||||   |||         |||   :::|||||||||  |||||||||   |||
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
                :::     ::::::     :::|||||||||      :::   :::
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
             ||||||||||||||||||   |||||||||   ||||||                ||||||
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
             ||||||   :::     |||:::|||   |||     |||||||||::::::
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
                :::         ||||||||||||||||||||   ::::::   :::          |||
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
             ||||||                ||||||:::         :::   |||||||||   |||
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
```

```
             ||||||   :::     |||         |||||||||||||||||  |||||||||
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
              |||||||||   ::::::   :::         ||||||||||||||||||||
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
                |||   |||||||||      :::||||||          ||||||||||||||||||
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
                  ::::::|||||||||    |||         ::::::|||||||||||||||||
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
              |||  ||||||||||||                    :::|||||:::   ||||||
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
              :::         :::|||   :::|||||||||:::                |||   |||
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
              ||||||||||||||||||:::         :::       |||||::: |||||||||
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
               ||||||||||||||||||   |||||||||:::                |||   |||||
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
              |||||||              :::|||   |||         |||||||||:::
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
              |||  :::   ||||||||||||||||||        :::   ||||||||||||   |
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
              ||   :::         :::||||||||||||||||||   ::::::      |||
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
                ||||||       |||   :::|||||||||||||||||||||       |||   ||
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
              |||||||   :::         |||         ||||||||||||||||||  |||||||||
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
                |||||||||::::::   |||         :::|||   ||||||||||||
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
                |||         |||:::|||||||||   ||||||:::   |||   :::   |||
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
                  :::   |||||||||   |||         :::       ||||||||||||||
Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
              |||||      |||   :::         ||||||||||||:::         :::||||||
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
              |||||||||       |||   :::||||||||||||         |||:::|||
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
              |||||||||||||||||||||:::   :::   :::      |||||||||||| |||
```

```
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
              |||||||||||||||||||| ||||||     :::|||||||||:::|||:::
Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
              |||||||||||||||||||||||   :::             |||   |||
Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
              ||||||        :::   :::         |||   |||||||||   |||||||||
Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
              |||            :::|||:::|||||||||||||||||||        :::   |||
Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
              |||||||||    |||      |||||||||||||||||||||      :::||
Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 632

Qy       2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
              |   :::   |||||||||   |||:::::|||  ||||||||||||||||||   :::|
Db        632 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 652

Qy       2836 AA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
              ||            |||||||||||| ||||||                |||||||||||
Db        652 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 669

Qy       2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
              ||||||      ||| :::   |||||||||   :::   |||   ::::::   |||||
Db        670 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 689

Qy       2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
              |||||||||||:::    |||   |||||||||   :::     :::|||      |
Db        689 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 709

Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
              ||||||||||||||||||:::        |||      ||||||||||||   :::
Db        709 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 728

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
              |||||||||||||||||   |||||||||:::::      :::   |||:::
Db        729 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 748

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
              |||||||||   ||||||||||||||||||                |||     ||||||
Db        749 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 768

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
              |||   |||        |||||||||||||||| |||       :::         |
Db        769 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 788

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
              ||   :::|||   |||      ||||||||||||||||||   |||||||||
Db        788 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 805

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
                |||      |||           |||||||||:::           |||   :::
Db        806 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 825

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
                |||||||||      :::           :::|||||||||||||||||
Db        825 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 845

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
                ||| |||||||||                           ||||||||||||
Db        845 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 864

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
              ||||||:::   |||::::::         |||||||| :::   |||   ||||||||
Db        865 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 884

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
              |||||||||||||    :::|||   ||||||||                :::   :::|
Db        884 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 904

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
              |||||||||||||||||       |||   :::   :::   |||   :::||:::||||
Db        904 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 924

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
              ||    |||||||||||||||||||:::|||   :::       ||||||
Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
              |||||||||:::            ||||||||||||                ||||||
Db        944 HisAlaHisLys------------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
              |||||||||   |||                |||||||||||||||||       :::|||
Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
              |||||||||         |||      |||   ||||||||||||||||||
Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996
```

RESULT 11
AAU33357
ID AAU33357 standard; protein; 1095 AA.
XX
AC AAU33357;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human breast cancer protein B726P fusion protein #1.
XX
KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy; immunogen.
XX
OS Homo sapiens.
XX
PN WO200179286-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US012164.
XX
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX
DR WPI; 2001-611721/70.
DR N-PSDB; AAS47421.
XX
PT Breast Tumor Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer.
XX
PS Claim 22; Page 292-295; 297pp; English.
XX
CC The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and

CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
XX
SQ Sequence 1095 AA;

Alignment Scores:

Pred. No.:	5.81e-95	Length:	1095
Score:	914.50	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.91%	Indels:	145
DB:	4	Gaps:	83

US-09-602-362E-22 (1-4115) x AAU33357 (1-1095)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db         94 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 113

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         114 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 133

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         134 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 153

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         154 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProProSerGluSerLysGlnGlu 173

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         174 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 192

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db         193 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 212

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db         213 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 232

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db         233 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 250

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db         251 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 270

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db         270 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 290

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db         290 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 310

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db         310 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737

Db         330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db         350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 369

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db         370 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 389

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db         389 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 409

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db         409 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 429

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db         429 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 449

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db         449 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 469

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db         469 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 488

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db         488 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 508

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db         508 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 528

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db         528 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 548

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db         548 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 568

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db         568 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 587

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db         588 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 606

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db         606 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 626

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db         626 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 645

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db         646 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 665

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db         666 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 685

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
```

```
Db        686 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 705

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
Db        706 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 725

Qy       2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
Db        725 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 745

Qy       2836 AA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
Db        745 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 762

Qy       2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
Db        763 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 782

Qy       2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
Db        782 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 802

Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
Db        802 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 821

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
Db        822 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 841

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
Db        842 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 861

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
Db        862 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 881

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
Db        881 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 898

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
Db        899 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 918

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
Db        918 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 938

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
Db        938 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 957

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
Db        958 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 977

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
Db        977 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 997

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
Db        997 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 1017

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
Db       1017 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 1036

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
Db       1037 HisAlaHisLys----------LysAlaAspAsnLysSerLysIleThrIleAspIle 1052

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
Db       1053 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 1072

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
Db       1072 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 1089
```

RESULT 12
ABG78924
ID ABG78924 standard; protein; 1095 AA.
XX
AC ABG78924;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human breast tumour polypeptide #15.
XX
KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN US2002085998-A1.
XX
PD 04-JUL-2002.
XX
PF 13-APR-2001; 2001US-00834759.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI Henderson RA;
XX
DR WPI; 2002-635657/68.
DR N-PSDB; ABS64022.
XX
PT Novel breast cancer polynucleotides and polypeptides encoded by the
PT polynucleotides, useful for detecting the presence of breast cancer in a
PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX
PS Disclosure; Page 223-225; 247pp; English.
XX
CC The invention relates to an isolated breast tumour polynucleotide and the
CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC detecting the presence of breast cancer in a patient, and in
CC pharmaceutical compositions for treating breast cancer. The sequences are
CC useful for stimulating an immune response in a patient and can therefore
CC be used in production of vaccines. The sequences are also useful for
CC detecting the presence of a cancer in a patient, by obtaining a
CC biological sample from the patient, contacting the biological sample with
CC a composition of the invention and detecting the amount of polynucleotide
CC that hybridizes to the sample. This sequence represents a human breast
CC tumour polypeptide of the invention
XX
SQ Sequence 1095 AA;

Alignment Scores:

Pred. No.:	5.81e-95	Length:	1095
Score:	914.50	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.91%	Indels:	145

DB: 5 Gaps: 83

US-09-602-362E-22 (1-4115) x ABG78924 (1-1095)

```
Db        782 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 802

Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
              ||||||||||||||||||||:::         |||       ||||||||||||   :::
Db        802 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 821

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
              ||||||||||||||||||  |||||||||::::::      :::   |||:::
Db        822 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 841

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
              |||||||||  ||||||||||||||||||          |||      ||||||
Db        842 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 861

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
              |||   |||       ||||||||||||||||| |||         :::         |
Db        862 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 881

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
              ||   :::|||   |||      ||||||||||||||||||   |||||||||
Db        881 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 898

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
                |||       |||            |||||||||:::          |||   :::
Db        899 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 918

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
                |||||||||       :::       :::||||||||||||||||||
Db        918 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 938

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
                ||| |||||||||                                ||||||||||||
Db        938 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 957

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
              ||||||:::   |||::::::         |||||||| :::   |||   ||||||||
Db        958 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 977

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
              ||||||||||||    :::|||   |||||||||                :::   :::|
Db        977 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 997

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
              ||||||||||||||||      |||  :::  :::   |||  :::|||:::||||
Db        997 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 1017

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
              ||    ||||||||||||||||||:::|||   :::      ||||||
Db       1017 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 1036

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
              |||||||||:::            ||||||||||                  ||||||
Db       1037 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 1052

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
              |||||||||   |||           ||||||||||||||||||       :::|||
Db       1053 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 1072

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
               |||||||||          |||     |||   ||||||||||||||||
Db       1072 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 1089

RESULT 13
ABJ37747
ID   ABJ37747 standard; protein; 1095 AA.
XX
AC   ABJ37747;
XX
DT   15-MAY-2003  (first entry)
XX
DE   Human tumour-related protein - SEQ ID No 493.
XX
KW   Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW   tumour; breast cancer; cancer; immune response stimulation.
XX
OS   Homo sapiens.
XX
PN   WO200283956-A1.
XX
PD   24-OCT-2002.
XX
PF   15-APR-2002; 2002WO-US012378.
XX
PR   13-APR-2001; 2001US-00834759.
PR   07-DEC-2001; 2001US-00007805.
PR   13-FEB-2002; 2002US-00076622.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Houghton RL,  Sleath PR,  Persing DH,  Jiang Y,  Dillon DC;
PI   Mitcham JL,  Xu J,  Harlocker SL,  Hepler WT,  Henderson RA,  Fanger GR;
PI   Vedvick TS,  Mcneill PD,  Durham M;
XX
DR   WPI; 2003-103376/09.
XX
PT   New polypeptide and polynucleotide useful for stimulating and/or
PT   expanding T cells specific for a tumor protein and treating breast
PT   cancer.
XX
PS   Disclosure; Page 310-312; 375pp; English.
XX
CC   The invention comprises a method of stimulating and/or expanding T cells
CC   specific for a tumour protein. The invention further comprises human
CC   nucleic acids and proteins that are associated with tumours (e.g. breast
CC   cancer). The method and sequences of the invention are useful for
CC   stimulating and/or expanding T cells specific for a tumour protein,
CC   detecting the presence of cancer, stimulating an immune response in a
CC   patient and treating breast cancer. The present amino acid sequence
CC   represents a human tumour-related protein
XX
SQ   Sequence 1095 AA;

Alignment Scores:
Pred. No.:              5.81e-95        Length:         1095
Score:                  914.50          Matches:        506
Percent Similarity:     59.77%          Conservative:   112
Best Local Similarity:  48.94%          Mismatches:     285
Query Match:            12.91%          Indels:         145
DB:                     6               Gaps:           83

US-09-602-362E-22 (1-4115) x ABJ37747 (1-1095)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
              |||||||||   |||:::      |||       |||||||||   ||||||||||||||
Db         94 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 113

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
              |||  ||||||||||||||:::   |||   |||   ||||||   ||||||||||||
Db        114 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 133

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
              |||||||||   |||       |||   :::|||||||||   |||||||||   |||
Db        134 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 153

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
                  :::      ::::::      :::|||||||||          :::   :::
Db        154 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProProSerGluSerLysGlnGlu 173

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
              |||||||||||||||||   |||||||||   ||||||                ||||||
Db        174 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 192

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
              ||||||   :::      |||:::|||   |||      |||||||||::::::
Db        193 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 212
```

```
Db        918 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 938

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
                 ||| |||||||||                              ||||||||||||
Db        938 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 957

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
              ||||||:::   |||:::::         |||||||| :::   |||   ||||||||
Db        958 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 977

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
              ||||||||||||    :::|||   |||||||||               :::   :::|
Db        977 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 997

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
              ||||||||||||||||      |||   :::   :::   |||   :::|||:::||||
Db        997 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 1017

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
              ||    |||||||||||||||||||:::|||   :::      ||||||
Db       1017 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 1036

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
              |||||||||:::            ||||||||||                  ||||||
Db       1037 HisAlaHisLys------------LysAlaAspAsnLysSerLysIleThrIleAspIle 1052

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
              |||||||||   |||                |||||||||||||||||      :::|||
Db       1053 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 1072

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
               |||||||||         |||      |||   ||||||||||||||||||
Db       1072 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 1089
```

RESULT 14
ABJ37789
ID ABJ37789 standard; protein; 1239 AA.
XX
AC ABJ37789;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human tumour-related protein - SEQ ID No 577.
XX
KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX
OS Homo sapiens.
XX
PN WO200283956-A1.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, Mcneill PD, Durham M;
XX
DR WPI; 2003-103376/09.
XX
PT New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
PS Example 13; Page 359-363; 375pp; English.
XX

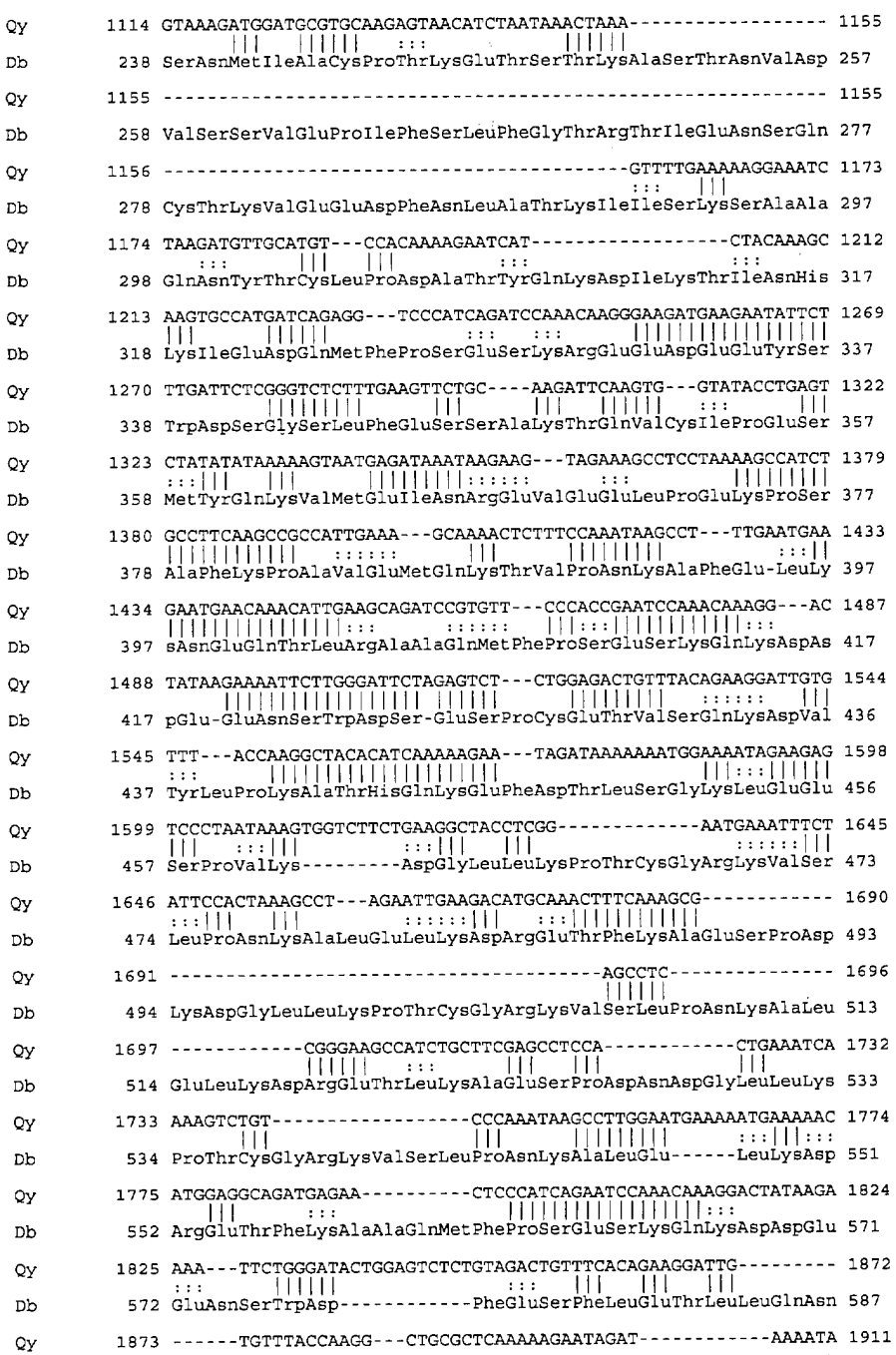
CC The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 1239 AA;

Alignment Scores:

Pred. No.:	3.76e-60	Length:	1239
Score:	612.50	Matches:	458
Percent Similarity:	45.75%	Conservative:	166
Best Local Similarity:	33.58%	Mismatches:	385
Query Match:	8.64%	Indels:	363
DB:	6	Gaps:	90

US-09-602-362E-22 (1-4115) x ABJ37789 (1-1239)

```
Qy        380 CATGGTGCAGTATCGAAGTGCCAA---CAAGGCTGCCTCACACCACTTTTACTACCAT-- 434
              :::|||||||||  :::   |||  :::      ||||||||||||||||||
Db          8 TyrGlyAlaValIleGluValGlnAsnLysAlaSerLeuThrProLeuLeuLeu-AlaIl 27

Qy        435 -AACGAAAGAAGTGAGCAATTGTGGAATTTT---GCTGATAAAAATGCAAATGCGAATGC 490
                  :::||||||:::|||          |||         ||||||||||||||||||
Db         27 eGlnLysArgSerLysGlnThrValGluPheLeuLeuThrLysAsnAlaAsnAlaAsnAl 47

Qy        491 GTTA---ATAAGTTAAATGCACACCCTCATGCTGCTGTATG---TCTGGATCATCAGAGA 544
              |          ||| |||||||||               :::      ||||||||||||
Db         47 aPheAsnGluSer-LysCysThrAlaLeuMetLeuAlaIleCysGluGlySerSerGluI 67

Qy        545 TAGTTGCATGCTTCTTAGCAAAA---TGTGACGTCTTTGTGCAG---ATATAGTGGAGTA 598
              |||||         |||   :::      |||||||||   :::   |||    |||:::
Db         67 leValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaGluAspIle-HisGlyIle 86

Qy        599 ACTGCAGAA---CATATGCTGTTACTGTGGATTTCTCACATTCATAACAAATTATGAATA 655
              |||||||||   :::                :::::::::||||||   :::|||
Db         87 ThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIleHisGlnGlnLeuLeuGlu 106

Qy        656 TAT---ACGAAAATTATCAAAAATCATCAAATACCAATCAGAAGGAACTCTGCAGGAAAC 712
              :::      |||:::   ||||||
Db        107 HisIleArgLysLeuProLysAsn---------------------------------- 114

Qy        713 CTGATGAGGCTGCACCCTGGCGGAAAGACACCTGACACGCTGAAAGCTGGTGGAAAAACA 772
                           |||          |||:::           |||      |||
Db        115 --------------ProGlnAsnThrAsnProGluGlyThrSerThrGly------Thr 127

Qy        773 CCTGATGAGGCTGCACCCT---GGTGGAAAGACACCTGACAGGCTGAAAGCT---GGTGG 826
              |||||||||||||| |||         :::|||||||||      ::::::
Db        128 ProAspGluAlaAla-ProLeuAlaGluArgThrProAspThrAlaGluSerLeuLeuGl 147

Qy        827 AAAAACACCTGATGAG---GCTGCATCTTGGTGGAGGAACATCTGACAAATTCAAT---G 880
               ||||||||||||||   |||     |||       |||||||||   |||         :
Db        147 uLysThrProAspGluAlaAla-ArgLeuValGluGlyThrSerAlaLysIleGlnCysL 167

Qy        881 TTGGAGAAAGGACATCTGGAAAGTTCGAACGTCAGC---AGAAAAACACCTAGGAAA--- 934
              ::|||:::   |||||||||||||||||||::: |||      :::||||||||||||
Db        167 euGlyLysAlaThrSerGlyLysPheGluGln-SerThrGluGluThrProArgLysIle 186

Qy        935 TTACGATCCTGCAAAAAAACATCTGAGAAATTTACGGG-CCAGCAAAGGAAGACCTAGAA 993
              ||||||      |||:::||||||||||||||:::   ||||||||||||
Db        187 LeuArgProThrLysGluThrSerGluLysPheSerTrpProAlaLysGlu--------- 203

Qy        994 GATCGCAGGGAGAAAAAGAAGACACACCTAGGGAAATATGAGTCCCGAAAAGAAACACTG 1053
                  |||   :::|||   |||             :::|||   ::::::   ::::::
Db        204 ------ArgSerArgLysIleThr-----------TrpGluGluLysGluThrSerVal 217

Qy       1054 AGAAATTACGTGGGCAGAAAAGGAAGACCTAGGAAGACGCATGGGAGAAAAAGAAACCCT 1113
              :::                          |||   ||||||         ::::::
Db        218 LysThrGluCysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThr 237
```


Qy 1114 GTAAAGATGGATGCGTGCAAGAGTAACATCTAATAAACTAAA---------- 1155
Db 238 SerAsnMetIleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAsp 257
Qy 1155 ---------- 1155
Db 258 ValSerSerValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGln 277
Qy 1156 ----------GTTTTGAAAAAGGAAATC 1173
Db 278 CysThrLysValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAla 297
Qy 1174 TAAGATGTTGCATGT---CCACAAAAGAATCAT----------CTACAAAGC 1212
Db 298 GlnAsnTyrThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHis 317
Qy 1213 AAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGGGAAGATGAAGAATATTCT 1269
Db 318 LysIleGluAspGlnMetPheProSerGluSerLysArgGluGluAspGluGluTyrSer 337
Qy 1270 TTGATTCTCGGGTCTCTTTGAAGTTCTGC----AAGATTCAAGTG---GTATACCTGAGT 1322
Db 338 TrpAspSerGlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSer 357
Qy 1323 CTATATATAAAAAGTAATGAGATAAATAAGAAG---TAGAAAGCCTCCTAAAAGCCATCT 1379
Db 358 MetTyrGlnLysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSer 377
Qy 1380 GCCTTCAAGCCGCCATTGAAA---GCAAAACTCTTTCCAAATAAGCCT---TTGAATGAA 1433
Db 378 AlaPheLysProAlaValGluMetGlnLysThrValProAsnLysAlaPheGlu-LeuLy 397
Qy 1434 GAATGAACAAACATTGAAGCAGATCCGTGTT---CCCACCGAATCCAAACAAAGG---AC 1487
Db 397 sAsnGluGlnThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAs 417
Qy 1488 TATAAGAAAATTCTTGGGATTCTAGAGTCT---CTGGAGACTGTTTACAGAAGGATTGTG 1544
Db 417 pGlu-GluAsnSerTrpAspSer-GluSerProCysGluThrValSerGlnLysAspVal 436
Qy 1545 TTT---ACCAAGGCTACACATCAAAAAGAA---TAGATAAAAAAATGGAAAATAGAAGAG 1598
Db 437 TyrLeuProLysAlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluGlu 456
Qy 1599 TCCCTAATAAAGTGGTCTTCTGAAGGCTACCTCGG----------AATGAAATTTCT 1645
Db 457 SerProValLys---------AspGlyLeuLeuLysProThrCysGlyArgLysValSer 473
Qy 1646 ATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTCAAAGCG---------- 1690
Db 474 LeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaGluSerProAsp 493
Qy 1691 ----------AGCCTC---------- 1696
Db 494 LysAspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeu 513
Qy 1697 ----------CGGGAAGCCATCTGCTTCGAGCCTCCA----------CTGAAATCA 1732
Db 514 GluLeuLysAspArgGluThrLeuLysAlaGluSerProAspAsnAspGlyLeuLeuLys 533
Qy 1733 AAAGTCTGT----------CCCAAATAAGCCTTGGAATGAAAAATGAAAAAC 1774
Db 534 ProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu------LeuLysAsp 551
Qy 1775 ATGGAGGCAGATGAGAA----------CTCCCATCAGAATCCAAACAAAGGACTATAAGA 1824
Db 552 ArgGluThrPheLysAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAspGlu 571
Qy 1825 AAA---TTCTGGGATACTGGAGTCTCTGTAGACTGTTTCACAGAAGGATTG--------- 1872
Db 572 GluAsnSerTrpAsp----------PheGluSerPheLeuGluThrLeuLeuGlnAsn 587
Qy 1873 ------TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGAT----------AAAATA 1911
Db 588 AspValCysLeuProLysAlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeu 607
Qy 1912 ATGGAAAATTAGAAGGGTCCCTGTTAAAGTGGTCTTCTGAGGCTAACT---GGGAATGAA 1968
Db 608 GluGluSer---------ProAspLysAspGlyLeuLeuLysProThrCysGlyMetLys 624
Qy 1969 AGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAAAGCAGGCCTC 2025
Db 625 IleSerLeuProAsnLysAla-LeuGluLeuLysAspArgGluThrPheLys------Al 642
Qy 2026 CCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAATCAAAAGTCTGTCCAAATAAACCTTG 2085
Db 642 aGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThrGlu-- 661
Qy 2086 GAATTAAGAATGAACAAACATTGAGGCAGAT----------GAGAACTCCCA 2127
Db 662 -AsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAlaTh 681
Qy 2128 TCAAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAG---------- 2173
Db 681 rLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGlnAs 701
Qy 2173 ---------- 2173
Db 701 pGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThrSe 721
Qy 2174 ----------TCTCTGTAGACTGTTTCCAGAAGGATGGTGTTTACCCAGGCTAC 2217
Db 721 rAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAlaTh 741
Qy 2218 ACATCAAAAAGAATAGATAAAA------TAATGGAAAATAGAAGAGTCTCTGATAATGAG 2271
Db 741 rTyrGlnLysGlu---IleLysThrThrAsnGlyLysIleGluGluSer---------- 756
Qy 2272 GTTTTCTGAAGGCTCCCTGAGAATGAAAGTTCTATTCCACTAAAGCCTTGAATTGATGGC 2331
Db 756 ---------- 756
Qy 2332 ATGCAAACTTTCAAAGCAGGCCTCCCGAGAGCCATCTGCTTCGAGCCTGCATTGAAATGA 2391
Db 757 -------ProGluLysProSerHisPheGluProAlaThrGluMet----------Gl 770
Qy 2392 AAAGTCTGTTCCAAATAAA------CCTTGGAATTAAGAATGAACAACATTGAGACAGAT 2445
Db 770 nAsnSerValProAsnLysGlyLeuGluTrpLysAsnLysGlnThr---------- 785
Qy 2446 CAGATTTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAATTCTGGGATTCTGAAGTCT 2505
Db 786 ----------LeuArgAlaAspSerThrThrLeuSerLysIleLeu-AspAlaLeuProS 802
Qy 2506 CCGTGGACTGTTTCACAGAAGGATGGTGTGTACCCAGGCTACACACAAAAAGAAAGGATA 2565
Db 802 erCysGlu---------- 804
Qy 2566 AAATAGTGGAAAATTAGAAGATTCACTAGCCTATCAAAATCTTGGTACAGTTCATCTTGT 2625
Db 804 ---------- 804
Qy 2626 GAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACACGTAC---AGGAAAATGGAACAAT 2682
Db 805 --ArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLysMetGluGlnM 824
Qy 2683 GAAAAAGAGTTTTGTGTACT---GAAAAAGAACTGTCAGAACAAAAGAAATAAATCACA- 2738
Db 824 etLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGluIleLysSerG 844
Qy 2739 --GTAGAGAACCAAAAGTT-AAATGGGAACAAGAG---TCTGCAGTGTAGATTGACTTAA 2792
Db 844 lnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArgLeuProLeuA 864
Qy 2793 ACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAAAATTAGGAAGAATT 2849

```
Db        864 snGlnGluGluGluLysArg-ArgAsnValAspIleLeuLysGluLysIle-ArgProGl 883

Qy       2850 AGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAAGTGAAACACAACTTGAAAGGCTCT-- 2907

Db        883 uGluGlnLeuArg---------LysLysLeuGluValLysHisGlnLeuGluGlnThrLe 900

Qy       2908 -CAGATACAAGAT---AAGAATTGAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA--- 2960

Db        900 uArgIleGlnAspIleGlu-LeuLysSerValThrSerAsnLeuAsnGln-ValSerHis 919

Qy       2961 ACTCATGAAATGAAA---ATTATTCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA 3017

Db        920 ThrHisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAla 939

Qy       3018 ---TGCAAAACTGGAATAGCCACACTA---AACACCAATCCAGGAAAAGGAAAATAAATC 3071

Db        940 MetLeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyr 959

Qy       3072 TTTGAGGACTTAAGATTTTAAAGA---AAAGATGCTGAACTTAG-ATGACCCTAAAACTG 3127

Db        960 PheGluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLeu 979

Qy       3128 AAAAGGAATCATTACTAAAAG---GGATCTCAATAT---GTGGGCAGCTAAAGTTCTGAT 3181

Db        980 LysGlnLysThrValThrLysArgAlaSerGlnTyrArgGlu-GlnLeuLysValLeuTh 999

Qy       3182 AGCTGAGAAACAATGCTCATTCTAAATTGAGGAAAAACAGACAAAGAAAA------CTAG 3235

Db        999 rAlaGlu-AsnThrMetLeuThrSerLysLeuLysGluLysGlnAspLysGluIleLeuG 1019

Qy       3236 AGGCAGAAATTGAATA-CACCATCCTGACTGGCTTCGCTGTACAAGC-----------C 3282

Db       1019 luThrGluIleGluSerHisHisPro---------ArgLeuAlaSerAlaLeuGlnAspH 1036

Qy       3283 ATGATCAATT-GTGACATCAAGAAAAAGTAAGAACCTGCTTCCACAT---TGAGGAGATG 3338

Db       1036 isAspGlnSerValThrSerArgLysAsnGlnGluLeuAlaPheHisSerAlaGlyAspA 1056

Qy       3339 CTGTTTGCAAAGAAA---ATGAATGTTGATGTG---AGAGTACGATATTAACAATGAG-- 3390

Db       1056 laPro-LeuGlnGlyIleMetAsnValAspValSerAsnThrIleTyr-AsnAsnGluVa 1075

Qy       3391 -TGCTCCATCACCACTTTCTGAGCTCAAAGGAAATCCAAAACCTAAAAA------TTATC 3443

Db       1075 lLeuHisGlnProLeuTyrGluAlaGlnArgLysSerLysSerProLysIleAsnLeuAs 1095

Qy       3444 TCAATTAGCAGGAGATGTCTAAGAGAAATACATTGGTTTCAGAACATCACAAAGAGACAA 3503

Db       1095 nTyrAlaGlyAspAspLeu-ArgGluAsnAlaLeuValSerGluHis-AlaGlnArgAsp 1114

Qy       3504 CGT------GAAAACAGTGTCAATGAAGGAAGCGAACACATGTATCAAAAC------GAC 3551

Db       1115 ArgCysGluThrGlnCysGlnMetLysLysAlaGluHisMetTyrGlnAsnGluGlnAsp 1134

Qy       3552 AAGATAATGGAACAAACACCTGAACAGCAGAGTCTCTAGTCAGAAATTATTTCAACTACA 3611

Db       1135 AsnValAspLysHisThrGluGlnGlnGluSerLeu-GluGlnLysLeuPheGlnLeuGl 1154

Qy       3612 AGCAAAAATTGTGGCTTCACAGCAATTAGTCATGCACATAGAAAGCTGACAACAAAAGCA 3671

Db       1154 uSerLysAsnArgTrpLeuArgGlnGlnLeuValTyrAlaHisLysLysValAsnLysSe 1174

Qy       3672 GATAACAATGATATTCATTTCTTGAGAGGAAATG---------CAACACATCTCCTAAAA 3722

Db       1174 rLysValThrIleAsnIleGlnPheProGluMetLysMetGlnArgHisLeu------Ly 1192

Qy       3723 GAGAAAAAGAGGAGATATT---AATTACAATACCATTTAAAAACCGTATATT---CAATA 3776

Db       1192 sGluLysAsnGluGluValPheAsnTyrGlyAsnHisLeuLysGluArgIleAspGlnTy 1212

Qy       3777 TGAAAAAGAGAAAGAGAAACAGAAACTC 3804

Db       1212 rGluLysGluLysAlaGluArgGluVal 1221
```

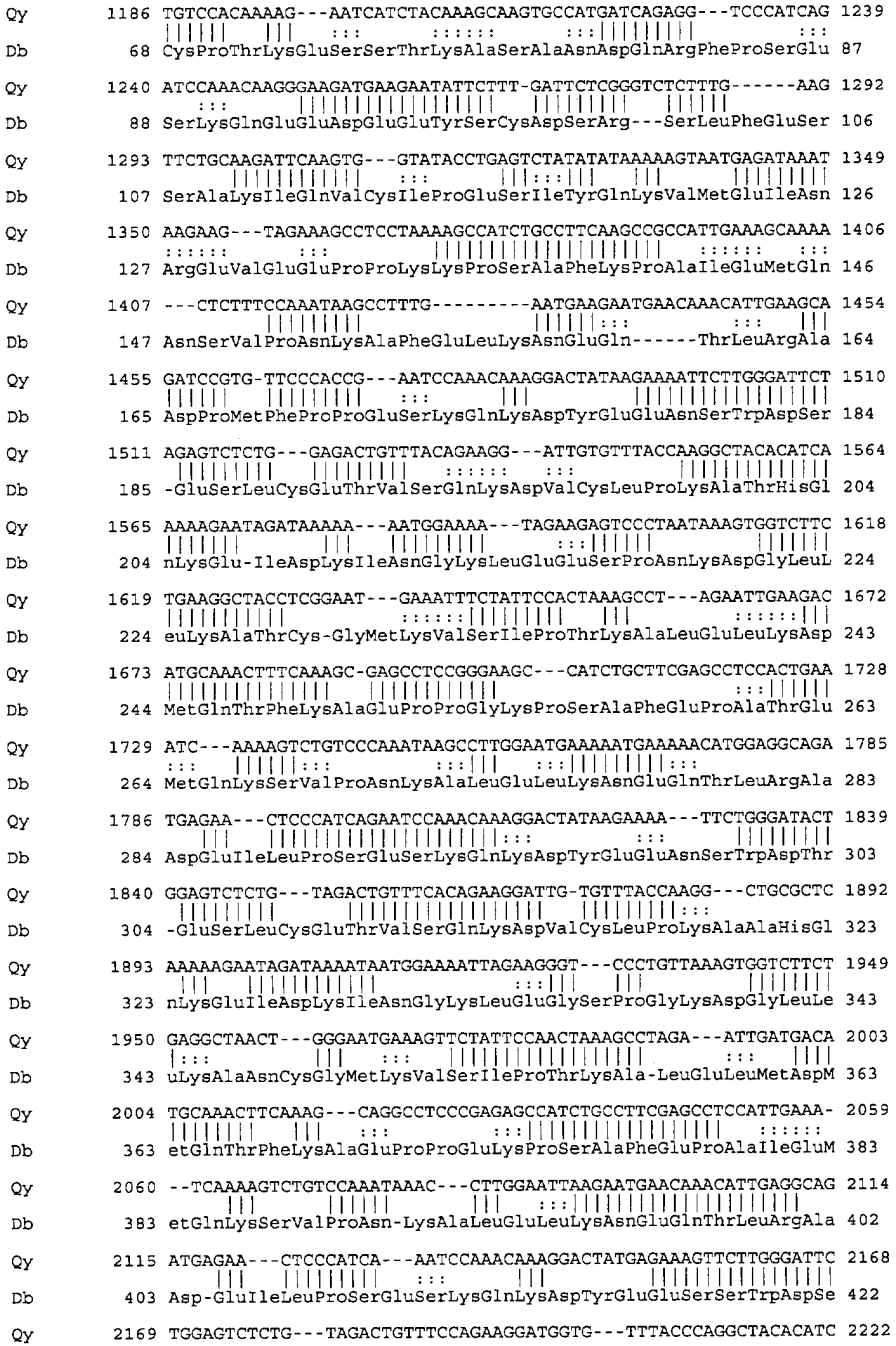
RESULT 15
ABJ37782
ID ABJ37782 standard; protein; 661 AA.
XX
AC ABJ37782;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human tumour-related protein - SEQ ID No 552.
XX
KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX
OS Homo sapiens.
XX
PN WO200283956-A1.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, Mcneill PD, Durham M;
XX
DR WPI; 2003-103376/09.
XX
PT New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
PS Example 8; Page 332-334; 375pp; English.
XX
CC The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 661 AA;

Alignment Scores:

Pred. No.:	1.62e-58	Length:	661
Score:	597.00	Matches:	315
Percent Similarity:	59.78%	Conservative:	73
Best Local Similarity:	48.54%	Mismatches:	191
Query Match:	8.43%	Indels:	79
DB:	6	Gaps:	52

US-09-602-362E-22 (1-4115) x ABJ37782 (1-661)

```
Qy       1018 CACCTAGGGAAATATGAGTCCCGAAAAGAAACACTGAGAAATTACGTG---GGCAGAAAA 1074

Db          8 HisValGlySerMetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLys 27

Qy       1075 GGAAGACCTAGGAAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT- 1129

Db         28 GlyArgProArgLysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysVal 47

Qy       1130 GCAAGAGTAACATCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCA 1185

Db         48 AlaArgValThrSerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAla 67
```



Search completed: July 15, 2004, 09:07:57
Job time : 308.082 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 08:47:52 ; Search time 34.9972 Seconds
(without alignments)
12140.450 Million cell updates/sec

Title: US-09-602-362E-22
Perfect score: 7086
Sequence: 1 ctagtctatacagcaacgac.....gagtggactccacctggaaa 4115

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093627_22040/app_query.fasta_1.10325
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09602362_@CGN_1_1_133_@runat_15072004_093627_22040 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	917	12.9	1002	4	US-09-620-405B-475	Sequence 475, App
2	917	12.9	1002	4	US-09-604-287A-475	Sequence 475, App
3	917	12.9	1002	4	US-09-834-759-475	Sequence 475, App
4	914.5	12.9	1095	4	US-09-620-405B-493	Sequence 493, App
5	914.5	12.9	1095	4	US-09-834-759-493	Sequence 493, App
6	583.5	8.2	650	4	US-09-620-405B-469	Sequence 469, App
7	583.5	8.2	650	4	US-09-433-826B-469	Sequence 469, App
8	583.5	8.2	650	4	US-09-604-287A-469	Sequence 469, App
9	583.5	8.2	650	4	US-09-834-759-469	Sequence 469, App
10	582	8.2	743	4	US-09-620-405B-494	Sequence 494, App
11	582	8.2	743	4	US-09-834-759-494	Sequence 494, App
12	421.5	5.9	445	4	US-09-620-405B-473	Sequence 473, App
13	421.5	5.9	445	4	US-09-433-826B-473	Sequence 473, App
14	421.5	5.9	445	4	US-09-604-287A-473	Sequence 473, App
15	421.5	5.9	445	4	US-09-834-759-473	Sequence 473, App
16	416.5	5.9	432	4	US-09-389-681-181	Sequence 181, App
17	416.5	5.9	432	4	US-09-620-405B-181	Sequence 181, App
18	416.5	5.9	432	4	US-09-339-338-181	Sequence 181, App
19	416.5	5.9	432	4	US-09-433-826B-181	Sequence 181, App
20	416.5	5.9	432	4	US-09-604-287A-181	Sequence 181, App
21	416.5	5.9	432	4	US-09-285-480-181	Sequence 181, App
22	416.5	5.9	432	4	US-09-834-759-181	Sequence 181, App
23	408.5	5.8	466	4	US-09-620-405B-472	Sequence 472, App
24	408.5	5.8	466	4	US-09-433-826B-472	Sequence 472, App
25	408.5	5.8	466	4	US-09-604-287A-472	Sequence 472, App
26	408.5	5.8	466	4	US-09-834-759-472	Sequence 472, App
27	285	4.0	410	4	US-09-620-405B-495	Sequence 495, App
28	285	4.0	410	4	US-09-834-759-495	Sequence 495, App
29	279	3.9	317	4	US-09-389-681-176	Sequence 176, App
30	279	3.9	317	4	US-09-620-405B-176	Sequence 176, App
31	279	3.9	317	4	US-09-339-338-176	Sequence 176, App
32	279	3.9	317	4	US-09-433-826B-176	Sequence 176, App
33	279	3.9	317	4	US-09-604-287A-176	Sequence 176, App
34	279	3.9	317	4	US-09-285-480-176	Sequence 176, App
35	279	3.9	317	4	US-09-834-759-176	Sequence 176, App
36	279	3.9	323	4	US-09-620-405B-483	Sequence 483, App
37	279	3.9	323	4	US-09-604-287A-483	Sequence 483, App
38	279	3.9	323	4	US-09-834-759-483	Sequence 483, App
39	208.5	2.9	228	4	US-09-620-405B-470	Sequence 470, App
40	208.5	2.9	228	4	US-09-433-826B-470	Sequence 470, App
41	208.5	2.9	228	4	US-09-604-287A-470	Sequence 470, App
42	208.5	2.9	228	4	US-09-834-759-470	Sequence 470, App
43	203	2.9	179	4	US-09-389-681-179	Sequence 179, App
44	203	2.9	179	4	US-09-620-405B-179	Sequence 179, App
45	203	2.9	179	4	US-09-339-338-179	Sequence 179, App

ALIGNMENTS

RESULT 1
US-09-620-405B-475
; Sequence 475, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-475

Alignment Scores:

Pred. No.:	9.7e-94	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	4	Gaps:	83

US-09-602-362E-22 (1-4115) x US-09-620-405B-475 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db          21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db          41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db          61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db          81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db         100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db         120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db         140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db         158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db         177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db         197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db         217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db         237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db         257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db         277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db         296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db         316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db         336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db         356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db         376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db         395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db         415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db         435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db         455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db         475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db         495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db         513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db         533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db         553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db         573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db         593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
Db         613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 632

Qy       2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
Db         632 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 652

Qy       2836 AA-----------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
Db         652 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 669

Qy       2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
Db         670 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 689

Qy       2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
Db         689 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 709
```

```
Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
Db        709 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 728

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
Db        729 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 748

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
Db        749 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 768

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
Db        769 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 788

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
Db        788 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 805

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
Db        806 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 825

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
Db        825 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 845

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
Db        845 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 864

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
Db        865 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 884

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
Db        884 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 904

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
Db        904 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 924

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
Db        944 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996
```

RESULT 2
US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

Alignment Scores:

Pred. No.:	9.7e-94	Length:	1002
Score:	917.00	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.94%	Indels:	145
DB:	4	Gaps:	83

US-09-602-362E-22 (1-4115) x US-09-604-287A-475 (1-1002)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236
```

```
Db        944 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
              |||||||||   |||           |||||||||||||||||||       :::|||
Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
               |||||||||          |||       |||   ||||||||||||||||||
Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996

RESULT 3
US-09-834-759-475
; Sequence 475, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 1002
;   TYPE: PRT
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(1002)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Alignment Scores:
Pred. No.:              9.7e-94       Length:          1002
Score:                  917.00        Matches:          506
Percent Similarity:     59.77%        Conservative:     112
Best Local Similarity:  48.94%        Mismatches:       285
Query Match:            12.94%        Indels:           145
DB:                     4             Gaps:              83

US-09-602-362E-22 (1-4115) x US-09-834-759-475 (1-1002)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
              |||||||||   |||:::      |||      |||||||||   ||||||||||||||||
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
              |||  ||||||||||||||:::   |||   |||   ||||||   ||||||||||||
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
              |||||||||   |||         |||   :::|||||||||   |||||||||   |||
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
                  :::      ::::::      :::|||||||||          :::   :::
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgPheProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
              ||||||||||||||||||   |||||||||   ||||||                ||||||
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
              ||||||   :::      |||:::|||   |||      |||||||||::::::
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
              :::                  ||||||||||||||||||||   ::::::   :::      |||
Db         120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
              ||||||                ||||||:::         :::   |||||||||  |||
Db         140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
              ||||||   :::      |||          ||||||||||||||||||| |||||||||
Db         158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
               ||||||||  ::::::   :::          ||||||||||||||||||||
Db         177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
                |||   |||||||||      :::||||||          ||||||||||||||||||
Db         197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
                  ::::::|||||||||   |||          ::::::||||||||||||||||||
Db         217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
              |||   ||||||||||||                        :::||||||:::   ||||||
Db         237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
              :::          :::|||   :::||||||||:::                 |||   |||
Db         257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
              ||||||||||||||||||:::          :::      ||||||::: |||||||||
Db         277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
               |||||||||||||||||   |||||||||:::                |||  |||||
Db         296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
              |||||||              :::|||   |||          |||||||||:::
Db         316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
              |||   :::   ||||||||||||||||||        :::   ||||||||||||   |
Db         336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
              ||   :::         :::|||||||||||||||||||   ::::::      |||
Db         356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysServ 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
              ||||||        |||   :::||||||||||||||||||||||||        |||   ||
Db         376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
              |||||||   :::      |||          |||||||||||||||||| |||||||||
Db         395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
                |||||||||::::::   |||          :::|||   |||||||||||
Db         415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
                |||      |||:::|||||||||   ||||||:::   |||   :::   |||
Db         435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
                  :::   |||||||||   |||          :::      ||||||||||||
Db         455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
```

```
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 632

Qy       2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
Db        632 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 652

Qy       2836 AA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
Db        652 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 669

Qy       2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
Db        670 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 689

Qy       2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
Db        689 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 709

Qy       2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
Db        709 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 728

Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
Db        729 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 748

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
Db        749 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 768

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
Db        769 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 788

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
Db        788 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 805

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
Db        806 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 825

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
```

```
Db        825 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 845

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
Db        845 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 864

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
Db        865 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 884

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
Db        884 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 904

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
Db        904 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 924

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
Db        924 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 943

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
Db        944 HisAlaHisLys------------LysAlaAspAsnLysSerLysIleThrIleAspIle 959

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
Db        960 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 979

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
Db        979 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 996
```

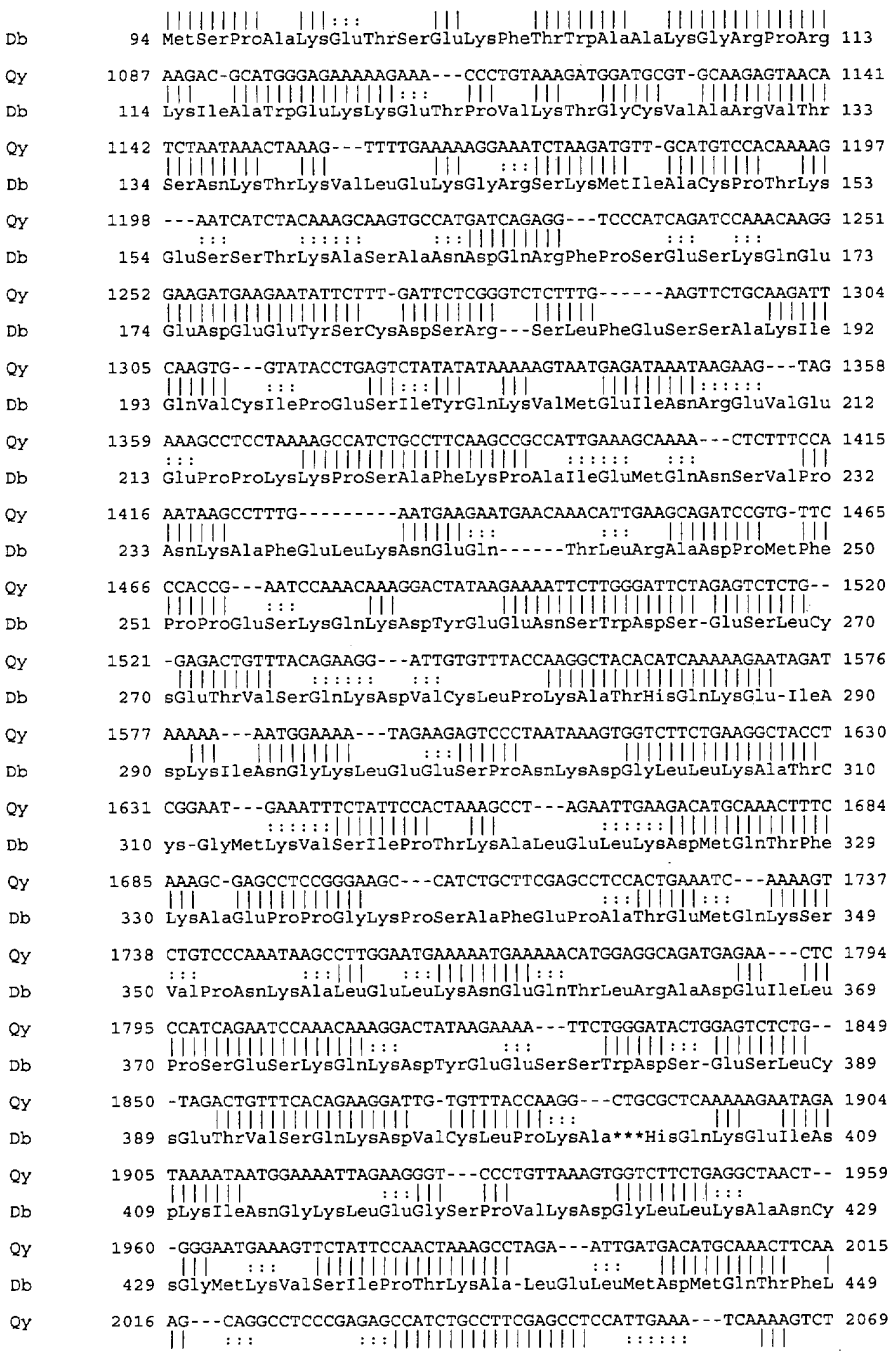
RESULT 4
US-09-620-405B-493
; Sequence 493, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-493

Alignment Scores:

Pred. No.:	1.94e-93	Length:	1095
Score:	914.50	Matches:	506
Percent Similarity:	59.77%	Conservative:	112
Best Local Similarity:	48.94%	Mismatches:	285
Query Match:	12.91%	Indels:	145
DB:	4	Gaps:	83

US-09-602-362E-22 (1-4115) x US-09-620-405B-493 (1-1095)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
```


Db 94 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 113
Qy 1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db 114 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 133
Qy 1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db 134 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 153
Qy 1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db 154 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 173
Qy 1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db 174 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 192
Qy 1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db 193 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 212
Qy 1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db 213 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 232
Qy 1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db 233 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 250
Qy 1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db 251 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 270
Qy 1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db 270 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 290
Qy 1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db 290 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 310
Qy 1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db 310 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329
Qy 1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db 330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349
Qy 1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db 350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 369
Qy 1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db 370 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 389
Qy 1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db 389 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 409
Qy 1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db 409 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 429
Qy 1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db 429 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 449
Qy 2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db 449 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 469
Qy 2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db 469 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 488
Qy 2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db 488 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 508
Qy 2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db 508 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 528
Qy 2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db 528 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 548
Qy 2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db 548 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 568
Qy 2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db 568 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 587
Qy 2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db 588 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 606
Qy 2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db 606 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 626
Qy 2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db 626 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 645
Qy 2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db 646 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 665
Qy 2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db 666 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 685
Qy 2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db 686 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 705
Qy 2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
Db 706 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 725
Qy 2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
Db 725 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 745
Qy 2836 AA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
Db 745 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 762
Qy 2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
Db 763 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 782
Qy 2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
Db 782 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 802
Qy 2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
Db 802 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 821

```
Qy       3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
Db        822 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 841

Qy       3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
Db        842 AlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGln 861

Qy       3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
Db        862 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 881

Qy       3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
Db        881 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 898

Qy       3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
Db        899 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 918

Qy       3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
Db        918 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 938

Qy       3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
Db        938 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 957

Qy       3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
Db        958 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 977

Qy       3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
Db        977 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 997

Qy       3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
Db        997 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 1017

Qy       3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
Db       1017 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 1036

Qy       3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
Db       1037 HisAlaHisLys------------LysAlaAspAsnLysSerLysIleThrIleAspIle 1052

Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
Db       1053 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 1072

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
Db       1072 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 1089

RESULT 5
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 493
;   LENGTH: 1095
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: variant
;   LOCATION: (1)...(1095)
;   OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Alignment Scores:
Pred. No.:              1.94e-93        Length:         1095
Score:                  914.50          Matches:        506
Percent Similarity:     59.77%          Conservative:   112
Best Local Similarity:  48.94%          Mismatches:     285
Query Match:            12.91%          Indels:         145
DB:                     4               Gaps:           83

US-09-602-362E-22 (1-4115) x US-09-834-759-493 (1-1095)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db         94 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 113

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db        114 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 133

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db        134 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 153

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db        154 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 173

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db        174 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 192

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        193 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 212

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        213 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 232

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        233 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 250

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        251 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 270

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        270 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 290

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        290 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 310

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        310 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
```


Db 330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349
Qy 1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db 350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 369
Qy 1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db 370 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 389
Qy 1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db 389 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 409
Qy 1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db 409 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 429
Qy 1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db 429 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 449
Qy 2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db 449 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 469
Qy 2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db 469 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 488
Qy 2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db 488 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 508
Qy 2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db 508 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 528
Qy 2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db 528 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 548
Qy 2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db 548 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 568
Qy 2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db 568 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 587
Qy 2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db 588 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 606
Qy 2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db 606 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 626
Qy 2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db 626 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 645
Qy 2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db 646 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 665
Qy 2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db 666 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 685
Qy 2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db 686 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 705
Qy 2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG---TCTGCAGT 2778
Db 706 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVa 725
Qy 2779 GTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATATATTAAAGAAAA 2835
Db 725 lArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIleLeuAsnGluL 745
Qy 2836 AA-----------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGAAAGAGTTAGAA 2883
Db 745 ysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------LysGluLeuGlu 762
Qy 2884 GTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATTGAAGGTGTAGA 2937
Db 763 ValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGl 782
Qy 2938 AAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTATTCTTACATGAA 2991
Db 782 uSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluA 802
Qy 2992 ATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCACACTA---AACA 3045
Db 802 snCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHis- 821
Qy 3046 CCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAAGA---AAAGAT 3101
Db 822 GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsn 841
Qy 3102 GCTGAACTTAG-ATGACCCTAAAACTGAAAAGGAATCATTACTAAAAG---GGATCTCAA 3157
Db 842 AlaGluLeuGlnMetThrLeuLysLeuLysGluSerLeuThrLysArgAlaSerGln 861
Qy 3158 TATGTGGGCAGCT---AAAGTTCTGATAGCTGAGAAACAATGCTCATTCTAAATTGAGGA 3214
Db 862 TyrSerGly-GlnLeuLysValLeuIleAlaGlu-AsnThrMetLeuThrSerLysLeuL 881
Qy 3215 AAAACAGACAAAGAAAA------CTAGAGGCAGAAATTGAATA-CACCATCCTGACTGGC 3267
Db 881 ysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro------- 898
Qy 3268 TTCGCTGTACAAGC------------CATGATCAATTGTGACATCAAGAAAAAGTAAGA- 3314
Db 899 --ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnG 918
Qy 3315 --ACCTGCTTCCACATTGAG---GAGATGCTGTTTGCAAAGAAAATGAATGTTGATGTG- 3368
Db 918 luProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValS 938
Qy 3369 --AGAGTACGATATTAACAATGAGTGCTCCATCACCACTTTC---TGAGCTCAAAGGAAA 3423
Db 938 erSerThrIleTyr-AsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLys 957
Qy 3424 TCCAAAACCTAAAAATTATCT---CAATTAGCAGGAGATGTCTAAGAGA---AATACATT 3477
Db 958 SerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAsp-AlaLeuArgGluAsnThrLe 977
Qy 3478 GGTTTCAGAACATCACAAAGAGA---CAACGTGAAAACAGTGTCAAT---GAAGGAAGCG 3531
Db 977 uValSerGluHis-AlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaG 997
Qy 3532 AACACATGTATCAAAAC------GACAAGATAATGGAACAAACACCTGAACAGCAGAGTC 3585
Db 997 luHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerL 1017
Qy 3586 TCTAGTCAGAAATTATTTCAACTACAAGCAAAAATTGTGGCTTCACAGCAATT----AGT 3641
Db 1017 eu-AspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuVal 1036
Qy 3642 CATGCACATAGAAAGCTGACAACAAAAGCAGATAACAAT--------------GATATT 3686
Db 1037 HisAlaHisLys-----------LysAlaAspAsnLysSerLysIleThrIleAspIle 1052

```
Qy       3687 CATTTCTTG---AGAGGAAATGCAACACATCTCCTAAAAGAGAAAAAGAGGAGATATT-- 3741
Db       1053 HisPheLeuGluArgLysMetGlnHisHisLeuLeuLysGluLysAsnGlu-GluIlePh 1072

Qy       3742 -AATTACAATACCATTTAAAAACCGTATATT---CAATATGAAAAAGAGAAA 3789
Db       1072 eAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLysGluLys 1089
```

RESULT 6
US-09-620-405B-469
; Sequence 469, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-469

Alignment Scores:

Pred. No.:	2.65e-56	Length:	650
Score:	583.50	Matches:	317
Percent Similarity:	60.47%	Conservative:	73
Best Local Similarity:	49.15%	Mismatches:	185
Query Match:	8.23%	Indels:	80
DB:	4	Gaps:	52

US-09-602-362E-22 (1-4115) x US-09-620-405B-469 (1-650)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysServ 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494
```

```
Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 628
```

RESULT 7
US-09-433-826B-469
; Sequence 469, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-433-826B-469

Alignment Scores:

Pred. No.:	2.65e-56	Length:	650
Score:	583.50	Matches:	317
Percent Similarity:	60.47%	Conservative:	73
Best Local Similarity:	49.15%	Mismatches:	185
Query Match:	8.23%	Indels:	80
DB:	4	Gaps:	52

US-09-602-362E-22 (1-4115) x US-09-433-826B-469 (1-650)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysServ 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
```

```
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179

Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234

Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289

Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344

Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398

Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450

Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505

Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560

Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614

Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670

Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724

Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770

Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 628

RESULT 8
US-09-604-287A-469
; Sequence 469, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C7
;  CURRENT APPLICATION NUMBER: US/09/604,287A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (310)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (429)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (522)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-469

Alignment Scores:
Pred. No.:                2.65e-56        Length:         650
Score:                    583.50          Matches:        317
Percent Similarity:       60.47%          Conservative:   73
Best Local Similarity:    49.15%          Mismatches:     185
Query Match:              8.23%           Indels:         80
DB:                       4               Gaps:           52

US-09-602-362E-22 (1-4115) x US-09-604-287A-469 (1-650)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086

Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141

Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197

Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251

Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304

Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358

Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415

Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465

Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520

Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576

Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630

Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684

Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737

Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256
```

```
Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 628

RESULT 9
US-09-834-759-469
; Sequence 469, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (310)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (429)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (522)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-469

Alignment Scores:
Pred. No.:              2.65e-56        Length:         650
Score:                  583.50          Matches:        317
Percent Similarity:     60.47%          Conservative:   73
Best Local Similarity:  49.15%          Mismatches:     185
Query Match:            8.23%           Indels:         80
DB:                     4               Gaps:           52

US-09-602-362E-22 (1-4115) x US-09-834-759-469 (1-650)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         61 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProProSerGluSerLysGlnGlu 80

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         81 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 99

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db        100 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 119
```

```
Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db        120 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 139

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db        140 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 157

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db        158 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 177

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db        177 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 197

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db        197 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 217

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db        217 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 236

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db        237 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 256

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db        257 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 276

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db        277 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 296

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db        296 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 316

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db        316 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 336

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db        336 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 356

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db        356 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 376

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db        376 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 395

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db        395 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 415

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db        415 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 435

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db        435 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 455

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db        455 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 475

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db        475 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 494

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db        495 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 513

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db        513 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 533

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db        533 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 552

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db        553 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 572

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db        573 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 592

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db        593 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 612

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db        613 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 628
```

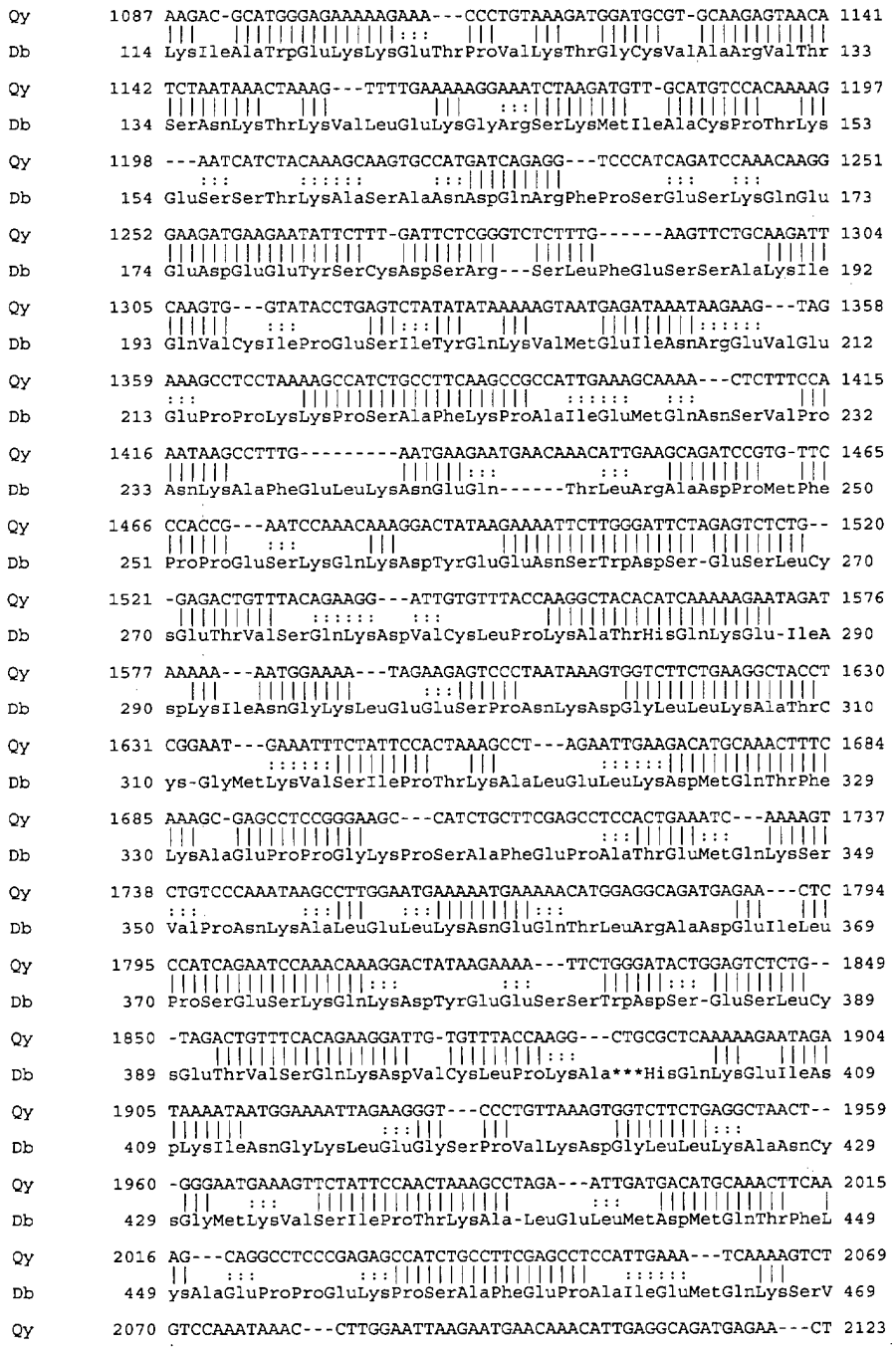
RESULT 10
US-09-620-405B-494
; Sequence 494, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-494

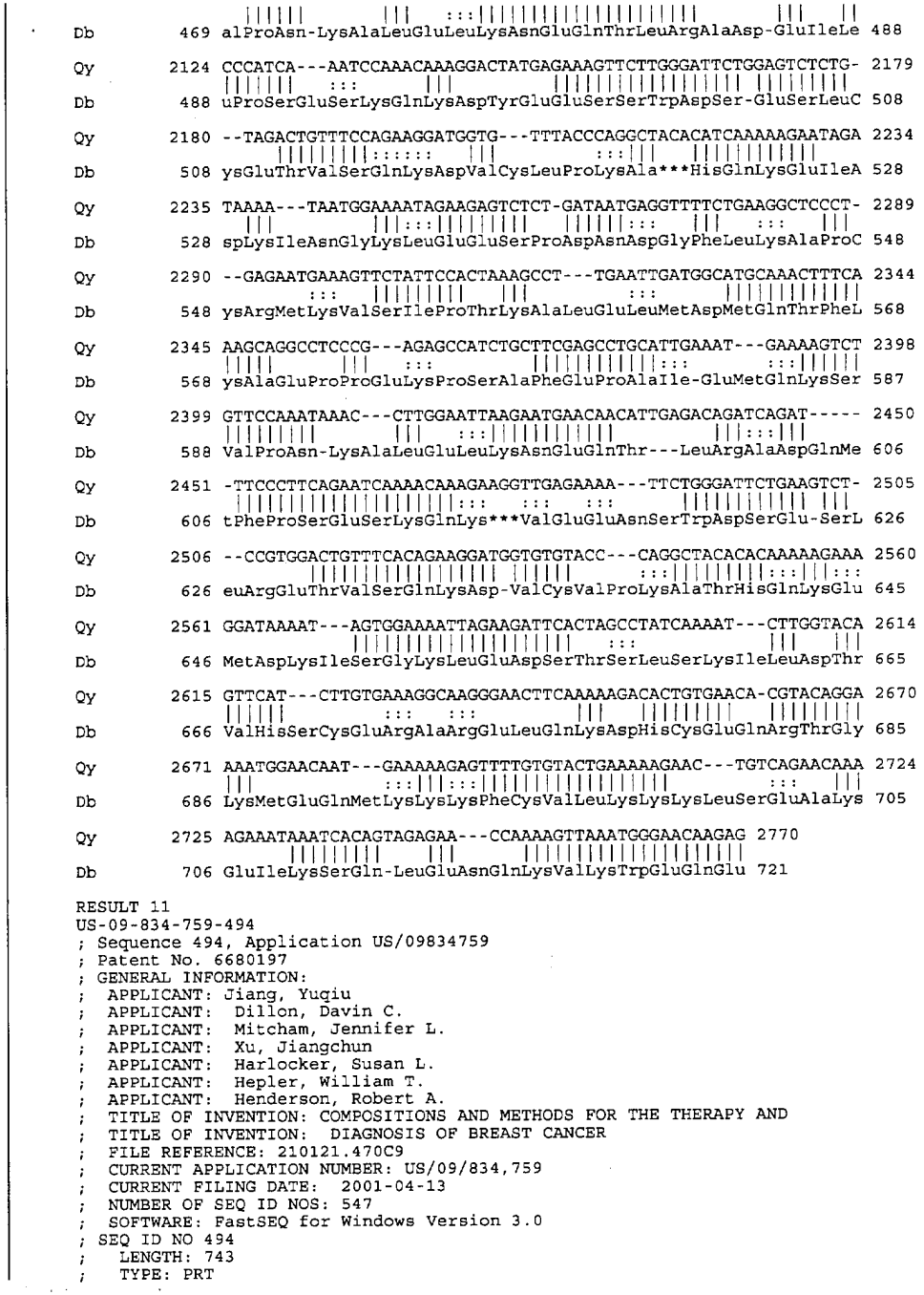
Alignment Scores:

Pred. No.:	4.2e-56	Length:	743
Score:	582.00	Matches:	317
Percent Similarity:	60.47%	Conservative:	73
Best Local Similarity:	49.15%	Mismatches:	185
Query Match:	8.21%	Indels:	80
DB:	4	Gaps:	52

US-09-602-362E-22 (1-4115) x US-09-620-405B-494 (1-743)

```
Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db         94 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 113
```





RESULT 11
US-09-834-759-494
; Sequence 494, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: variant
;   LOCATION: (1)...(743)
;   OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-494

Alignment Scores:
Pred. No.:                4.2e-56        Length:          743
Score:                    582.00         Matches:         317
Percent Similarity:       60.47%         Conservative:    73
Best Local Similarity:    49.15%         Mismatches:      185
Query Match:              8.21%          Indels:          80
DB:                       4              Gaps:            52

US-09-602-362E-22 (1-4115) x US-09-834-759-494 (1-743)

Qy       1031 ATGAGTCCCGAAAAGAAACACTGAGAA---ATTACGTGGGCAGA-AAAGGAAGACCTAGG 1086
Db         94 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 113

Qy       1087 AAGAC-GCATGGGAGAAAAAGAAA---CCCTGTAAAGATGGATGCGT-GCAAGAGTAACA 1141
Db         114 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 133

Qy       1142 TCTAATAAACTAAAG---TTTTGAAAAAGGAAATCTAAGATGTT-GCATGTCCACAAAAG 1197
Db         134 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 153

Qy       1198 ---AATCATCTACAAAGCAAGTGCCATGATCAGAGG---TCCCATCAGATCCAAACAAGG 1251
Db         154 GluSerSerThrLysAlaSerAlaAsnAspGlnArgProSerGluSerLysGlnGlu 173

Qy       1252 GAAGATGAAGAATATTCTTT-GATTCTCGGGTCTCTTTG------AAGTTCTGCAAGATT 1304
Db         174 GluAspGluGluTyrSerCysAspSerArg---SerLeuPheGluSerSerAlaLysIle 192

Qy       1305 CAAGTG---GTATACCTGAGTCTATATATATAAAAAGTAATGAGATAAATAAGAAG---TAG 1358
Db         193 GlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGluValGlu 212

Qy       1359 AAAGCCTCCTAAAAGCCATCTGCCTTCAAGCCGCCATTGAAAGCAAAA---CTCTTTCCA 1415
Db         213 GluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSerValPro 232

Qy       1416 AATAAGCCTTTG---------AATGAAGAATGAACAAACATTGAAGCAGATCCGTG-TTC 1465
Db         233 AsnLysAlaPheGluLeuLysAsnGluGln------ThrLeuArgAlaAspProMetPhe 250

Qy       1466 CCACCG---AATCCAAACAAAGGACTATAAGAAAATTCTTGGGATTCTAGAGTCTCTG-- 1520
Db         251 ProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSer-GluSerLeuCy 270

Qy       1521 -GAGACTGTTTACAGAAGG---ATTGTGTTTACCAAGGCTACACATCAAAAAGAATAGAT 1576
Db         270 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu-IleA 290

Qy       1577 AAAAA---AATGGAAAA---TAGAAGAGTCCCTAATAAAGTGGTCTTCTGAAGGCTACCT 1630
Db         290 spLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrC 310

Qy       1631 CGGAAT---GAAATTTCTATTCCACTAAAGCCT---AGAATTGAAGACATGCAAACTTTC 1684
Db         310 ys-GlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329

Qy       1685 AAAGC-GAGCCTCCGGGAAGC---CATCTGCTTCGAGCCTCCACTGAAATC---AAAAGT 1737
Db         330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349

Qy       1738 CTGTCCCAAATAAGCCTTGGAATGAAAAATGAAAAACATGGAGGCAGATGAGAA---CTC 1794
Db         350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeu 369

Qy       1795 CCATCAGAATCCAAACAAAGGACTATAAGAAAA---TTCTGGGATACTGGAGTCTCTG-- 1849
Db         370 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuCy 389

Qy       1850 -TAGACTGTTTCACAGAAGGATTG-TGTTTACCAAGG---CTGCGCTCAAAAAGAATAGA 1904
Db         389 sGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAs 409

Qy       1905 TAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTGAGGCTAACT-- 1959
Db         409 pLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCy 429

Qy       1960 -GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACATGCAAACTTCAA 2015
Db         429 sGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMetGlnThrPheL 449

Qy       2016 AG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA---TCAAAAGTCT 2069
Db         449 ysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerV 469

Qy       2070 GTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGATGAGAA---CT 2123
Db         469 alProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp-GluIleLe 488

Qy       2124 CCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGGAGTCTCTG- 2179
Db         488 uProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer-GluSerLeuC 508

Qy       2180 --TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCAAAAAGAATAGA 2234
Db         508 ysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleA 528

Qy       2235 TAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTGAAGGCTCCCT- 2289
Db         528 spLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProC 548

Qy       2290 --GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCATGCAAACTTTCA 2344
Db         548 ysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheL 568

Qy       2345 AAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT---GAAAAGTCT 2398
Db         568 ysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluMetGlnLysSer 587

Qy       2399 GTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGATCAGAT----- 2450
Db         588 ValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArgAlaAspGlnMe 606

Qy       2451 -TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGATTCTGAAGTCT- 2505
Db         606 tPheProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGlu-SerL 626

Qy       2506 --CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACACACAAAAAGAAA 2560
Db         626 euArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrHisGlnLysGlu 645

Qy       2561 GGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT---CTTGGTACA 2614
Db         646 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 665

Qy       2615 GTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAACA-CGTACAGGA 2670
Db         666 ValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGly 685

Qy       2671 AAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---TGTCAGAACAAA 2724
Db         686 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 705

Qy       2725 AGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db         706 GluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 721

RESULT 12
```

```
US-09-620-405B-473
; Sequence 473, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C8
;  CURRENT APPLICATION NUMBER: US/09/620,405B
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 495
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
;   LENGTH: 445
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-620-405B-473

Alignment Scores:
Pred. No.:                3.72e-38        Length:          445
Score:                    421.50          Matches:         226
Percent Similarity:       61.49%          Conservative:    55
Best Local Similarity:    49.45%          Mismatches:      119
Query Match:              5.95%           Indels:          60
DB:                       4               Gaps:            38

US-09-602-362E-22 (1-4115) x US-09-620-405B-473 (1-445)

Qy       1894 AAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTG 1950
                 |||   |||||||||||          :::|||   |||       |||||||||
Db          1 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 20

Qy       1951 AGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACAT 2004
                 :::         |||   :::   ||||||||||||||||||       :::   |||||
Db          21 LysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMe 40

Qy       2005 GCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA-- 2059
                 |||||||   |||   :::         :::||||||||||||||||||   ::::::
Db          40 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMe 60

Qy       2060 -TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGA 2115
                  |||       ||||||       |||   :::||||||||||||||||||||
Db          60 tGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaA 80

Qy       2116 TGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 2169
                 |||   |||||||||   :::       |||          ||||||||||||||||
Db          80 sp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 99

Qy       2170 GGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCA 2223
                 |||||||||       |||||||||::::::   |||          :::|||   |||||
Db         100 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 119

Qy       2224 AAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTG 2279
                 |||||||       |||          |||:::|||||||||   ||||||:::   |||
Db         119 nLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLe 139

Qy       2280 AAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCAT 2333
                  :::   |||       :::   |||||||||   |||          :::        ||
Db         139 uLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMe 159

Qy       2334 GCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT- 2389
                 |||||||||||||||       |||   :::         ||||||||||||:::
Db         159 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluM 179

Qy       2390 --GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGA 2444
                   ::||||||||||||||||       |||   :::||||||||||||         |||
Db         179 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArg 197

Qy       2445 TCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGAT 2495
                 :::|||      |||||||||||||||||||:::::::::   :::      ||||||
Db         198 AlaAspGlnMetPheProSerGluSerLysGlnLysLysValGluGluAsnSerTrpAsp 217

Qy       2496 TCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACAC 2549
                 ||||||  |||         |||||||||||||||||  ||||||       :::|||||||
Db         218 SerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrH 237

Qy       2550 ACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT- 2605
                 ||:::||:::            ||||||||||||||||||||   :::
Db         237 isGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysI 257

Qy       2606 --CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAAC 2660
                   |||   |||||||||         :::   :::         |||   |||||||||
Db         257 leLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluG 277

Qy       2661 A-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---T 2713
                  ||||||||||            :::|||:::||||||||||||||||||
Db         277 lnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuS 297

Qy       2714 GTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
                   :::   |||       |||||||||   |||       ||||||||||||||||||||
Db         297 erGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 316

Qy       2771 ---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATAT 2824
                     :::|||   :::   |||||||||   |||::::::||| ||||||||||||||
Db         317 LeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIl 336

Qy       2825 ATTAAAGAAAAAA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGA 2872
                 ||||   :::|||           ||||||||||| |||||                |
Db         336 eLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------L 354

Qy       2873 AAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATT 2926
                 ||||||||||||||||       ||| :::   |||||||||   :::   |||   :::
Db         354 ysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeu 373

Qy       2927 GAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTAT 2980
                 :::   |||||||||||||||||:::   |||   ||||||||   :::      ::
Db         374 LysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLe 393

Qy       2981 TCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCAC 3037
                 :|||       |||||||||||||||||||:::         |||       ||||||||
Db         393 uLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaTh 413

Qy       3038 ACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAA 3093
                 ||||   :::         |||||||||||||||||   |||||||||::::::      :
Db         413 rLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuL 433

Qy       3094 GA---AAAGATGCTGAACTTAGA 3113
                 ::   |||:::|||||||||:::
Db         433 ysGluLysAsnAlaGluLeuGln 440

RESULT 13
US-09-433-826B-473
; Sequence 473, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION:  COMPOSITIONS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.470C4
;  CURRENT APPLICATION NUMBER: US/09/433,826B
;  CURRENT FILING DATE:  1999-11-03
;  NUMBER OF SEQ ID NOS:  474
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 473
;   LENGTH: 445
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

US-09-433-826B-473

Alignment Scores:

Pred. No.:	3.72e-38	Length:	445
Score:	421.50	Matches:	226
Percent Similarity:	61.49%	Conservative:	55
Best Local Similarity:	49.45%	Mismatches:	119
Query Match:	5.95%	Indels:	60
DB:	4	Gaps:	38

US-09-602-362E-22 (1-4115) x US-09-433-826B-473 (1-445)

```
Qy       1894 AAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTG 1950
Db          1 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 20

Qy       1951 AGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACAT 2004
Db         21 LysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMe 40

Qy       2005 GCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA-- 2059
Db         40 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMe 60

Qy       2060 -TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGA 2115
Db         60 tGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaA 80

Qy       2116 TGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 2169
Db         80 sp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 99

Qy       2170 GGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCA 2223
Db        100 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 119

Qy       2224 AAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTG 2279
Db        119 nLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLe 139

Qy       2280 AAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCAT 2333
Db        139 uLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMe 159

Qy       2334 GCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT- 2389
Db        159 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluM 179

Qy       2390 --GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGA 2444
Db        179 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArg 197

Qy       2445 TCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGAT 2495
Db        198 AlaAspGlnMetPheProSerGluSerLysGlnLysLysValGluGluAsnSerTrpAsp 217

Qy       2496 TCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACAC 2549
Db        218 SerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrH 237

Qy       2550 ACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT- 2605
Db        237 isGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysI 257

Qy       2606 --CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAAC 2660
Db        257 leLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluG 277

Qy       2661 A-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---T 2713
Db        277 lnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuS 297

Qy       2714 GTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db        297 erGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 316

Qy       2771 ---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATAT 2824
Db        317 LeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIl 336

Qy       2825 ATTAAAGAAAAAA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGA 2872
Db        336 eLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------L 354

Qy       2873 AAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATT 2926
Db        354 ysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeu 373

Qy       2927 GAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTAT 2980
Db        374 LysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLe 393

Qy       2981 TCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCAC 3037
Db        393 uLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaTh 413

Qy       3038 ACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAA 3093
Db        413 rLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuL 433

Qy       3094 GA---AAAGATGCTGAACTTAGA 3113
Db        433 ysGluLysAsnAlaGluLeuGln 440
```

RESULT 14
US-09-604-287A-473
; Sequence 473, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-287A-473

Alignment Scores:

Pred. No.:	3.72e-38	Length:	445
Score:	421.50	Matches:	226
Percent Similarity:	61.49%	Conservative:	55
Best Local Similarity:	49.45%	Mismatches:	119
Query Match:	5.95%	Indels:	60
DB:	4	Gaps:	38

US-09-602-362E-22 (1-4115) x US-09-604-287A-473 (1-445)

```
Qy       1894 AAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTG 1950
Db          1 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 20

Qy       1951 AGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACAT 2004
Db         21 LysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMe 40
```

```
Qy       2005 GCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA-- 2059
Db         40 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMe 60

Qy       2060 -TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGA 2115
Db         60 tGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaA 80

Qy       2116 TGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 2169
Db         80 sp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 99

Qy       2170 GGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCA 2223
Db        100 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 119

Qy       2224 AAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTG 2279
Db        119 nLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLe 139

Qy       2280 AAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCAT 2333
Db        139 uLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMe 159

Qy       2334 GCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT- 2389
Db        159 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluM 179

Qy       2390 --GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGA 2444
Db        179 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArg 197

Qy       2445 TCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGAT 2495
Db        198 AlaAspGlnMetPheProSerGluSerLysGlnLysLysValGluGluAsnSerTrpAsp 217

Qy       2496 TCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACAC 2549
Db        218 SerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrH 237

Qy       2550 ACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT- 2605
Db        237 isGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysI 257

Qy       2606 --CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAAC 2660
Db        257 leLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluG 277

Qy       2661 A-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---T 2713
Db        277 lnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuS 297

Qy       2714 GTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db        297 erGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 316

Qy       2771 ---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATAT 2824
Db        317 LeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIl 336

Qy       2825 ATTAAAGAAAAAA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGA 2872
Db        336 eLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------L 354

Qy       2873 AAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATT 2926
Db        354 ysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeu 373

Qy       2927 GAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTAT 2980
Db        374 LysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLe 393

Qy       2981 TCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCAC 3037

Db        393 uLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaTh 413

Qy       3038 ACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAA 3093
Db        413 rLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuL 433

Qy       3094 GA---AAAGATGCTGAACTTAGA 3113
Db        433 ysGluLysAsnAlaGluLeuGln 440

RESULT 15
US-09-834-759-473
; Sequence 473, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
;   LENGTH: 445
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-759-473

Alignment Scores:
Pred. No.:              3.72e-38        Length:         445
Score:                  421.50          Matches:        226
Percent Similarity:     61.49%          Conservative:   55
Best Local Similarity:  49.45%          Mismatches:     119
Query Match:            5.95%           Indels:         60
DB:                     4               Gaps:           38

US-09-602-362E-22 (1-4115) x US-09-834-759-473 (1-445)

Qy       1894 AAAAGAATAGATAAAATAATGGAAAATTAGAAGGGT---CCCTGTTAAAGTGGTCTTCTG 1950
Db          1 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 20

Qy       1951 AGGCTAACT---GGGAATGAAAGTTCTATTCCAACTAAAGCCTAGA---ATTGATGACAT 2004
Db         21 LysAlaAsnCysGlyMetLysValSerIleProThrLysAla-LeuGluLeuMetAspMe 40

Qy       2005 GCAAACTTCAAAG---CAGGCCTCCCGAGAGCCATCTGCCTTCGAGCCTCCATTGAAA-- 2059
Db         40 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMe 60

Qy       2060 -TCAAAAGTCTGTCCAAATAAAC---CTTGGAATTAAGAATGAACAAACATTGAGGCAGA 2115
Db         60 tGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaA 80

Qy       2116 TGAGAA---CTCCCATCA---AATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 2169
Db         80 sp-GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 99

Qy       2170 GGAGTCTCTG---TAGACTGTTTCCAGAAGGATGGTG---TTTACCCAGGCTACACATCA 2223
Db        100 -GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGl 119

Qy       2224 AAAAGAATAGATAAAA---TAATGGAAAATAGAAGAGTCTCT-GATAATGAGGTTTTCTG 2279
Db        119 nLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLe 139
```

```
Qy       2280 AAGGCTCCCT---GAGAATGAAAGTTCTATTCCACTAAAGCCT---TGAATTGATGGCAT 2333
Db        139 uLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMe 159

Qy       2334 GCAAACTTTCAAAGCAGGCCTCCCG---AGAGCCATCTGCTTCGAGCCTGCATTGAAAT- 2389
Db        159 tGlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIle-GluM 179

Qy       2390 --GAAAAGTCTGTTCCAAATAAAC---CTTGGAATTAAGAATGAACAACATTGAGACAGA 2444
Db        179 etGlnLysSerValProAsn-LysAlaLeuGluLeuLysAsnGluGlnThr---LeuArg 197

Qy       2445 TCAGAT------TTCCCTTCAGAATCAAAACAAAGAAGGTTGAGAAAA---TTCTGGGAT 2495
Db        198 AlaAspGlnMetPheProSerGluSerLysGlnLysLysValGluGluAsnSerTrpAsp 217

Qy       2496 TCTGAAGTCT---CCGTGGACTGTTTCACAGAAGGATGGTGTGTACC---CAGGCTACAC 2549
Db        218 SerGlu-SerLeuArgGluThrValSerGlnLysAsp-ValCysValProLysAlaThrH 237

Qy       2550 ACAAAAAGAAAGGATAAAAT---AGTGGAAAATTAGAAGATTCACTAGCCTATCAAAAT- 2605
Db        237 isGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysI 257

Qy       2606 --CTTGGTACAGTTCAT---CTTGTGAAAGGCAAGGGAACTTCAAAAAGACACTGTGAAC 2660
Db        257 leLeuAspThrValHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluG 277

Qy       2661 A-CGTACAGGAAAATGGAACAAT---GAAAAAGAGTTTTGTGTACTGAAAAAGAAC---T 2713
Db        277 lnArgThrGlyLysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuS 297

Qy       2714 GTCAGAACAAAAGAAATAAATCACAGTAGAGAA---CCAAAAGTTAAATGGGAACAAGAG 2770
Db        297 erGluAlaLysGluIleLysSerGln-LeuGluAsnGlnLysValLysTrpGluGlnGlu 316

Qy       2771 ---TCTGCAGTGTAGATTGACTTAAACCAA---GAAAAGAGAAGAGAGAAATGCCGATAT 2824
Db        317 LeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLysArg-ArgAsnAlaAspIl 336

Qy       2825 ATTAAAGAAAAAA------------TTAGGAAGAATTAGAAGAATCGAGAGCAGCATAGA 2872
Db        336 eLeuAsnGluLysIleArgGluGluLeuGlyArgIle-GluGluGlnHisArg------L 354

Qy       2873 AAGAGTTAGAAGTGAAACACAACTTGAAAG---GCTCTCAGATACAAG---ATAAGAATT 2926
Db        354 ysGluLeuGluValLysGlnGlnLeu-GluGlnAlaLeuArgIleGlnAspIleGluLeu 373

Qy       2927 GAAGGTGTAGAAAGTAATTTGAATAGGTTTCTCA---ACTCATGAAATGAAA---ATTAT 2980
Db        374 LysSerValGluSerAsnLeuAsnGln-ValSerHisThrHisGluAsnGluAsnTyrLe 393

Qy       2981 TCTTACATGAAATTGCATGTTGAAAAAGGAATTGCCA---TGCAAAACTGGAATAGCCAC 3037
Db        393 uLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaTh 413

Qy       3038 ACTA---AACACCAATCCAGGAAAAGGAAAATAAATC-TTTGAGGACTTAAGATTTTAAA 3093
Db        413 rLeuLysHis-GlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuL 433

Qy       3094 GA---AAAGATGCTGAACTTAGA 3113
Db        433 ysGluLysAsnAlaGluLeuGln 440
```

Search completed: July 15, 2004, 09:28:02
Job time : 112.997 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 08:47:52 ; Search time 31.2381 Seconds
(without alignments)
12140.450 Million cell updates/sec

Title: US-09-602-362E-26
Perfect score: 6448
Sequence: 1 caagagcttggcgatacaga.....tgacttctaaattgaaggaa 3673

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093627_22040/app_query.fasta_1.10325
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09602362_@CGN_1_1_133_@runat_15072004_093627_22040 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2372.5	36.8	1095	4	US-09-620-405B-493	Sequence 493, App
2	2372.5	36.8	1095	4	US-09-834-759-493	Sequence 493, App
3	2367	36.7	1002	4	US-09-620-405B-475	Sequence 475, App
4	2367	36.7	1002	4	US-09-604-287A-475	Sequence 475, App
5	2367	36.7	1002	4	US-09-834-759-475	Sequence 475, App
6	1775	27.5	743	4	US-09-620-405B-494	Sequence 494, App
7	1775	27.5	743	4	US-09-834-759-494	Sequence 494, App
8	1769.5	27.4	650	4	US-09-620-405B-469	Sequence 469, App
9	1769.5	27.4	650	4	US-09-433-826B-469	Sequence 469, App
10	1769.5	27.4	650	4	US-09-604-287A-469	Sequence 469, App
11	1769.5	27.4	650	4	US-09-834-759-469	Sequence 469, App
12	1414.5	21.9	445	4	US-09-620-405B-473	Sequence 473, App
13	1414.5	21.9	445	4	US-09-433-826B-473	Sequence 473, App
14	1414.5	21.9	445	4	US-09-604-287A-473	Sequence 473, App
15	1414.5	21.9	445	4	US-09-834-759-473	Sequence 473, App
16	1384	21.5	466	4	US-09-620-405B-472	Sequence 472, App
17	1384	21.5	466	4	US-09-433-826B-472	Sequence 472, App
18	1384	21.5	466	4	US-09-604-287A-472	Sequence 472, App
19	1384	21.5	466	4	US-09-834-759-472	Sequence 472, App
20	884.5	13.7	432	4	US-09-389-681-181	Sequence 181, App
21	884.5	13.7	432	4	US-09-620-405B-181	Sequence 181, App
22	884.5	13.7	432	4	US-09-339-338-181	Sequence 181, App
23	884.5	13.7	432	4	US-09-433-826B-181	Sequence 181, App
24	884.5	13.7	432	4	US-09-604-287A-181	Sequence 181, App
25	884.5	13.7	432	4	US-09-285-480-181	Sequence 181, App
26	884.5	13.7	432	4	US-09-834-759-181	Sequence 181, App
27	798.5	12.4	228	4	US-09-620-405B-470	Sequence 470, App
28	798.5	12.4	228	4	US-09-433-826B-470	Sequence 470, App
29	798.5	12.4	228	4	US-09-604-287A-470	Sequence 470, App
30	798.5	12.4	228	4	US-09-834-759-470	Sequence 470, App
31	700.5	10.9	1719	4	US-09-439-313-378	Sequence 378, App
32	700.5	10.9	1719	4	US-09-352-616A-378	Sequence 378, App
33	700.5	10.9	1719	4	US-09-636-215-378	Sequence 378, App
34	700.5	10.9	1719	4	US-09-685-166A-378	Sequence 378, App
35	586.5	9.1	671	4	US-09-439-313-380	Sequence 380, App
36	586.5	9.1	671	4	US-09-352-616A-380	Sequence 380, App
37	586.5	9.1	671	4	US-09-289-198-306	Sequence 306, App
38	586.5	9.1	671	4	US-09-636-215-380	Sequence 380, App
39	586.5	9.1	671	4	US-09-685-166A-380	Sequence 380, App
40	586.5	9.1	671	4	US-09-429-755-306	Sequence 306, App
41	583.5	9.0	656	4	US-09-439-313-379	Sequence 379, App
42	583.5	9.0	656	4	US-09-352-616A-379	Sequence 379, App
43	583.5	9.0	656	4	US-09-289-198-305	Sequence 305, App
44	583.5	9.0	656	4	US-09-636-215-379	Sequence 379, App
45	583.5	9.0	656	4	US-09-685-166A-379	Sequence 379, App

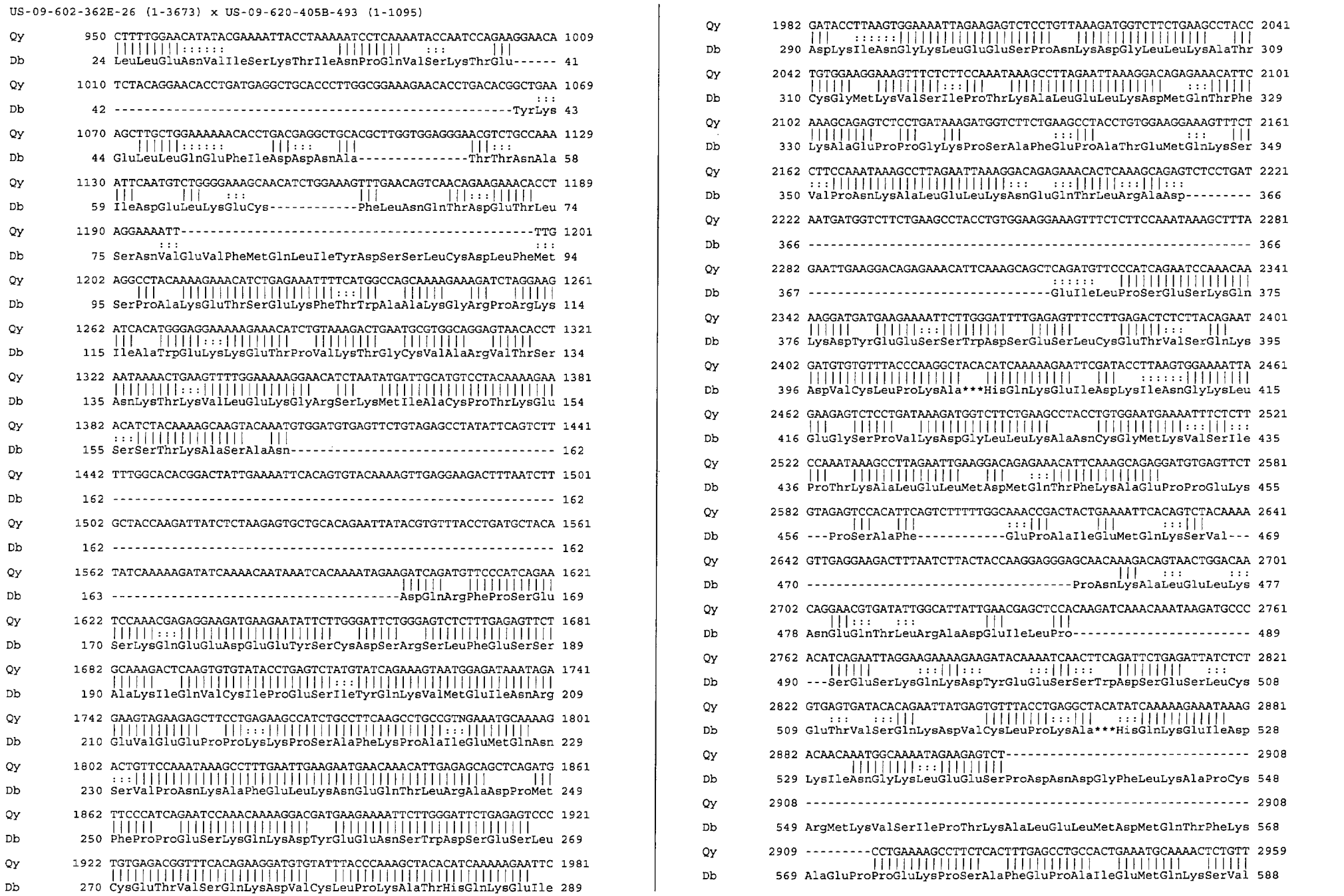
ALIGNMENTS

RESULT 1
US-09-620-405B-493
; Sequence 493, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-493

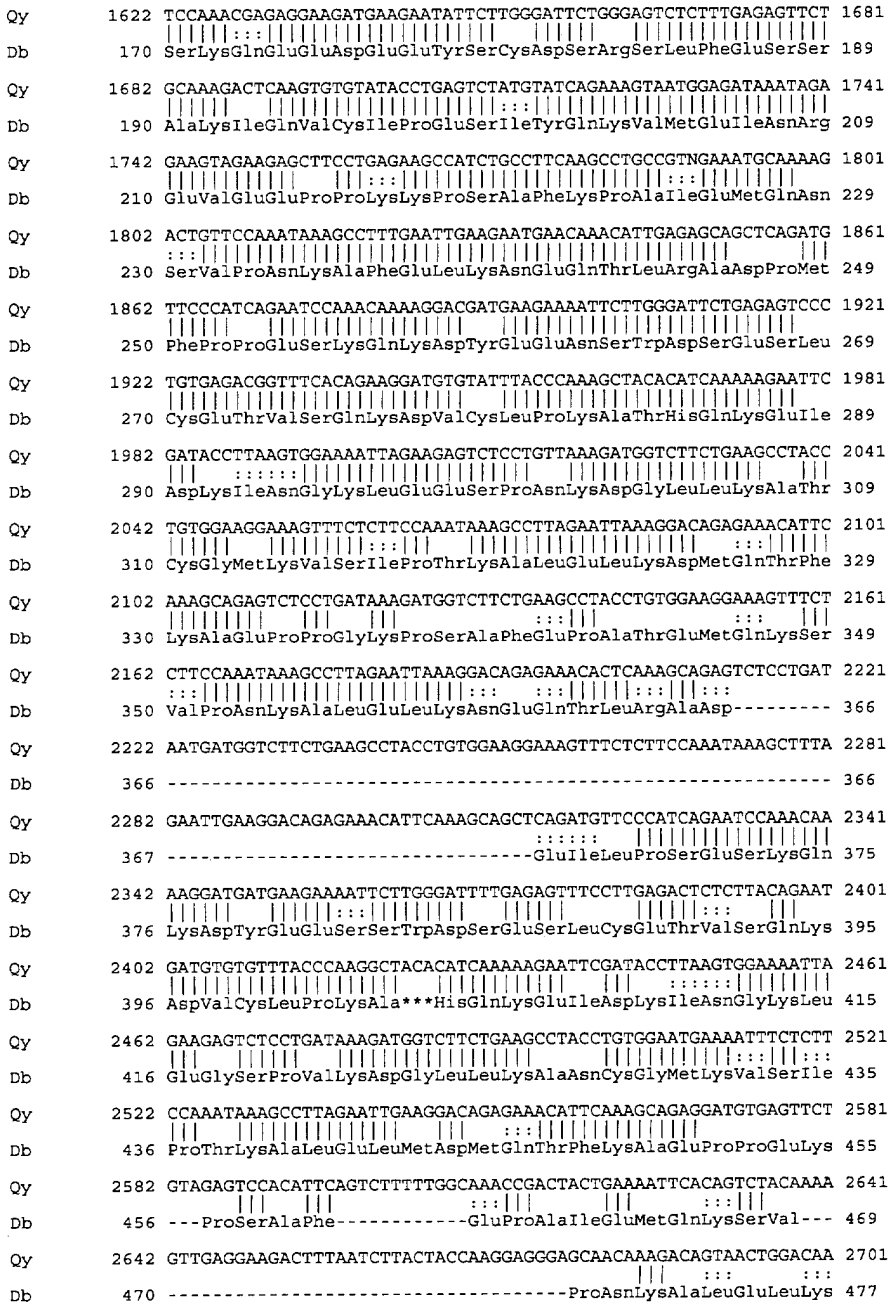
Alignment Scores:

Pred. No.:	6.47e-183	Length:	1095
Score:	2372.50	Matches:	529
Percent Similarity:	59.37%	Conservative:	73
Best Local Similarity:	52.17%	Mismatches:	151
Query Match:	36.79%	Indels:	261
DB:	4	Gaps:	13

US-09-602-362E-26 (1-3673) x US-09-620-405B-493 (1-1095)



```
Qy       2960 CCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db        589 ProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePro 608

Qy       3004 ---------------------------------------------------------- 3004
Db        609 SerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArgGlu 628

Qy       3004 ---------------------------------------------------------- 3004
Db        629 ThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLys 648

Qy       3005 ----------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCT 3046
Db        649 IleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSer 668

Qy       3047 TGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAA 3106
Db        669 CysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGlu 688

Qy       3107 CAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA 3166
Db        689 GlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLys 708

Qy       3167 TCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCT 3226
Db        709 SerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuThr 728

Qy       3227 TTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC 3286
Db        729 LeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGlu 748

Qy       3287 ---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAA 3331
Db        749 GluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGlu 768

Qy       3332 CAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTT 3391
Db        769 GlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVal 788

Qy       3392 TCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAA 3451
Db        789 SerHisThrHisGluAsnGluAsnTyrLeuHisGluAsnCysMetLeuLysLysGlu 808

Qy       3452 ATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAAT 3511
Db        809 IleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsn 828

Qy       3512 AAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTA 3571
Db        829 LysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeu 848

Qy       3572 AAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTT 3631
Db        849 LysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVal 868

Qy       3632 CTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db        869 LeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 882

RESULT 2
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 493
;   LENGTH: 1095
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: variant
;   LOCATION: (1)...(1095)
;   OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Alignment Scores:
Pred. No.:              6.47e-183      Length:          1095
Score:                  2372.50        Matches:         529
Percent Similarity:     59.37%         Conservative:    73
Best Local Similarity:  52.17%         Mismatches:      151
Query Match:            36.79%         Indels:          261
DB:                     4              Gaps:            13

US-09-602-362E-26 (1-3673) x US-09-834-759-493 (1-1095)

Qy        950 CTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACA 1009
Db         24 LeuLeuGluAsnValIleSerLysThrIleAsnProGlnValSerLysThrGlu------ 41

Qy       1010 TCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAA 1069
Db         42 ----------------------------------------------------TyrLys 43

Qy       1070 AGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAA 1129
Db         44 GluLeuGluGlnGluPheIleAspAspAsnAla--------------ThrThrAsnAla 58

Qy       1130 ATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCT 1189
Db         59 IleAspGluLeuLysGluCys-----------PheLeuAsnGlnThrAspGluThrLeu 74

Qy       1190 AGGAAAATT----------------------------------------------TTG 1201
Db         75 SerAsnValGluValPheMetGlnLeuIleTyrAspSerSerLeuCysAspLeuPheMet 94

Qy       1202 AGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAG 1261
Db         95 SerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLys 114

Qy       1262 ATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCT 1321
Db        115 IleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSer 134

Qy       1322 AATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAA 1381
Db        135 AsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGlu 154

Qy       1382 ACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTT 1441
Db        155 SerSerThrLysAlaSerAlaAsn---------------------------------- 162

Qy       1442 TTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTT 1501
Db        162 ---------------------------------------------------------- 162

Qy       1502 GCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACA 1561
Db        162 ---------------------------------------------------------- 162

Qy       1562 TATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAA 1621
Db        163 -----------------------------------AspGlnArgPheProSerGlu 169
```


Qy 1622 TCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCT 1681
Db 170 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSer 189
Qy 1682 GCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGA 1741
Db 190 AlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArg 209
Qy 1742 GAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAG 1801
Db 210 GluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsn 229
Qy 1802 ACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATG 1861
Db 230 SerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMet 249
Qy 1862 TTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCC 1921
Db 250 PheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLeu 269
Qy 1922 TGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTC 1981
Db 270 CysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIle 289
Qy 1982 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACC 2041
Db 290 AspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThr 309
Qy 2042 TGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTC 2101
Db 310 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329
Qy 2102 AAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT 2161
Db 330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349
Qy 2162 CTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGAT 2221
Db 350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp--------- 366
Qy 2222 AATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTA 2281
Db 366 ---------- 366
Qy 2282 GAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAA 2341
Db 367 ----------GluIleLeuProSerGluSerLysGln 375
Qy 2342 AAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAAT 2401
Db 376 LysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGlnLys 395
Qy 2402 GATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTA 2461
Db 396 AspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLysLeu 415
Qy 2462 GAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTT 2521
Db 416 GluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIle 435
Qy 2522 CCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCT 2581
Db 436 ProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys 455
Qy 2582 GTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAA 2641
Db 456 ---ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal--- 469
Qy 2642 GTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAA 2701
Db 470 ----------ProAsnLysAlaLeuGluLeuLys 477
Qy 2702 CAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCC 2761
Db 478 AsnGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 489
Qy 2762 ACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCT 2821
Db 490 ---SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 508
Qy 2822 GTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAG 2881
Db 509 GluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAsp 528
Qy 2882 ACAACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 529 LysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCys 548
Qy 2908 ---------- 2908
Db 549 ArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 568
Qy 2909 ---------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTT 2959
Db 569 AlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 588
Qy 2960 CCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 589 ProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePro 608
Qy 3004 ---------- 3004
Db 609 SerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArgGlu 628
Qy 3004 ---------- 3004
Db 629 ThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLys 648
Qy 3005 ----------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCT 3046
Db 649 IleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSer 668
Qy 3047 TGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAA 3106
Db 669 CysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGlu 688
Qy 3107 CAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA 3166
Db 689 GlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLys 708
Qy 3167 TCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCT 3226
Db 709 SerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuThr 728
Qy 3227 TTAAATCAAGAAGAAGAAGAAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC 3286
Db 729 LeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGlu 748
Qy 3287 ----------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAA 3331
Db 749 GluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGlu 768
Qy 3332 CAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTT 3391
Db 769 GlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVal 788
Qy 3392 TCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAA 3451
Db 789 SerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGlu 808
Qy 3452 ATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAAT 3511
Db 809 IleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsn 828
Qy 3512 AAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTA 3571

```
Db        829 LysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeu 848

Qy       3572 AAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTT 3631

Db        849 LysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVal 868

Qy       3632 CTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673

Db        869 LeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 882

RESULT 3
US-09-620-405B-475
; Sequence 475, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C8
;  CURRENT APPLICATION NUMBER: US/09/620,405B
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 495
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 1002
;   TYPE: PRT
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(1002)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-475

Alignment Scores:
Pred. No.:              1.74e-182       Length:         1002
Score:                  2367.00         Matches:        511
Percent Similarity:     62.62%          Conservative:   62
Best Local Similarity:  55.85%          Mismatches:     126
Query Match:            36.71%          Indels:         216
DB:                     4               Gaps:           9

US-09-602-362E-26 (1-3673) x US-09-620-405B-475 (1-1002)

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258

Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318

Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378

Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438

Db         61 GluSerSerThrLysAlaSerAlaAsn---------------------------------- 69

Qy       1439 CTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAAT 1498

Db         69 ------------------------------------------------------------ 69

Qy       1499 CTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCT 1558

Db         69 ------------------------------------------------------------ 69

Qy       1559 ACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCA 1618

Db         70 ----------------------------------------AspGlnArgPheProSer 75

Qy       1619 GAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGT 1678

Db         76 GluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSer 95

Qy       1679 TCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAAT 1738

Db         96 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 115

Qy       1739 AGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAA 1798

Db        116 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 135

Qy       1799 AAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAG 1858

Db        136 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 155

Qy       1859 ATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGT 1918

Db        156 MetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 175

Qy       1919 CCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAA 1978

Db        176 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195

Qy       1979 TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCT 2038

Db        196 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 215

Qy       2039 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2098

Db        216 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 235

Qy       2099 TTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2158

Db        236 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255

Qy       2159 TCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCT 2218

Db        256 SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp------ 273

Qy       2219 GATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCT 2278

Db        273 ------------------------------------------------------------ 273

Qy       2279 TTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAA 2338

Db        274 ----------------------------------GluIleLeuProSerGluSerLys 281

Qy       2339 CAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAG 2398

Db        282 GlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGln 301

Qy       2399 AATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAA 2458

Db        302 LysAspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLys 321

Qy       2459 TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCT 2518

Db        322 LeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 341

Qy       2519 CTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGT 2578

Db        342 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGlu 361

Qy       2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACA 2638

Db        362 Lys---ProSerAlaPhe-----------GluProAlaIleGluMetGlnLysSerVal 376

Qy       2639 AAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGA 2698
```

```
Db        377 ------------------------------------ProAsnLysAlaLeuGluLeu 383

Qy       2699 CAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATG 2758
Db        384 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro---------------------- 396

Qy       2759 CCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATC 2818
Db        397 ------SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 414

Qy       2819 TCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATA 2878
Db        415 CysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIle 434

Qy       2879 AAGACAACAAATGGCAAAATAGAAGAGTCT---------------------------- 2908
Db        435 AspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaPro 454

Qy       2908 ------------------------------------------------------------ 2908
Db        455 CysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 474

Qy       2909 ------------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCT 2956
Db        475 LysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 494

Qy       2957 GTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA------------ 3004
Db        495 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhe 514

Qy       3004 ------------------------------------------------------------ 3004
Db        515 ProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArg 534

Qy       3004 ------------------------------------------------------------ 3004
Db        535 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 554

Qy       3005 --------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCT 3043
Db        555 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 574

Qy       3044 TCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATG 3103
Db        575 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 594

Qy       3104 GAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATA 3163
Db        595 GluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 614

Qy       3164 AAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTG 3223
Db        615 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeu 634

Qy       3224 CCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGA 3283
Db        635 ThrLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArg 654

Qy       3284 CCC--------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTT 3328
Db        655 GluGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeu 674

Qy       3329 GAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAG 3388
Db        675 GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln 694

Qy       3389 GTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAG 3448
Db        695 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 714

Qy       3449 GAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAA 3508

Db        715 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGlu 734

Qy       3509 AATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACC 3568
Db        735 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 754

Qy       3569 CTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAA 3628
Db        755 LeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 774

Qy       3629 GTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db        775 ValLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 789
```

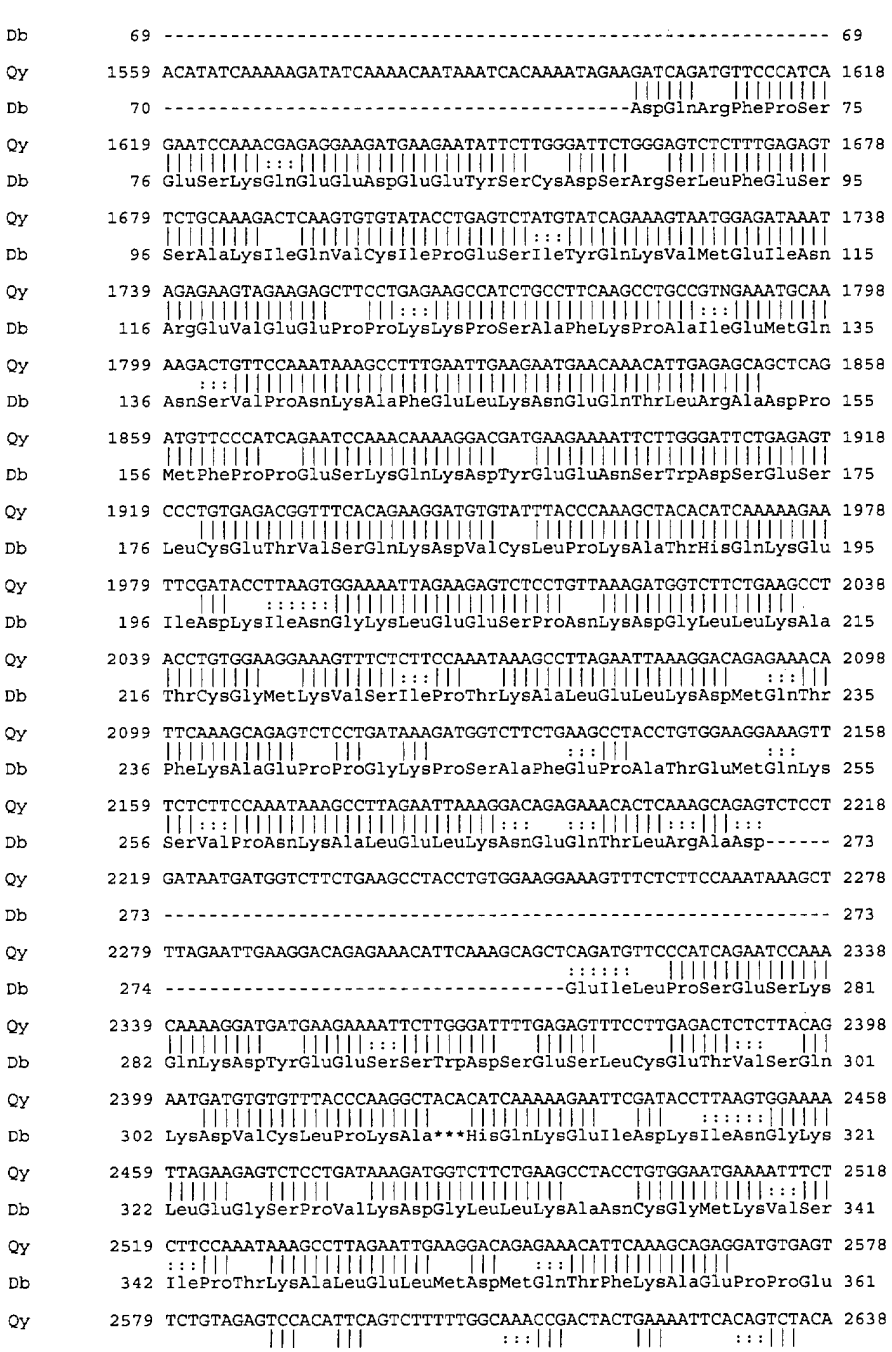
RESULT 4
US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

Alignment Scores:

Pred. No.:	1.74e-182	Length:	1002
Score:	2367.00	Matches:	511
Percent Similarity:	62.62%	Conservative:	62
Best Local Similarity:	55.85%	Mismatches:	126
Query Match:	36.71%	Indels:	216
DB:	4	Gaps:	9

US-09-602-362E-26 (1-3673) x US-09-604-287A-475 (1-1002)

```
Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438
Db         61 GluSerSerThrLysAlaSerAlaAsn---------------------------------- 69

Qy       1439 CTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAAT 1498
Db         69 ------------------------------------------------------------ 69

Qy       1499 CTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCT 1558
```


Db 69 ---------- 69
Qy 1559 ACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCA 1618
Db 70 ----------AspGlnArgPheProSer 75
Qy 1619 GAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGT 1678
Db 76 GluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSer 95
Qy 1679 TCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAAT 1738
Db 96 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 115
Qy 1739 AGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAA 1798
Db 116 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 135
Qy 1799 AAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAG 1858
Db 136 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 155
Qy 1859 ATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGT 1918
Db 156 MetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 175
Qy 1919 CCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAA 1978
Db 176 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195
Qy 1979 TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCT 2038
Db 196 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 215
Qy 2039 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2098
Db 216 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 235
Qy 2099 TTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2158
Db 236 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255
Qy 2159 TCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCT 2218
Db 256 SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp------ 273
Qy 2219 GATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCT 2278
Db 273 ---------- 273
Qy 2279 TTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAA 2338
Db 274 ----------GluIleLeuProSerGluSerLys 281
Qy 2339 CAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAG 2398
Db 282 GlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGln 301
Qy 2399 AATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAA 2458
Db 302 LysAspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLys 321
Qy 2459 TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCT 2518
Db 322 LeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 341
Qy 2519 CTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGT 2578
Db 342 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGlu 361
Qy 2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACA 2638
Db 362 Lys---ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal 376
Qy 2639 AAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGA 2698
Db 377 ----------ProAsnLysAlaLeuGluLeu 383
Qy 2699 CAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATG 2758
Db 384 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 396
Qy 2759 CCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATC 2818
Db 397 ------SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 414
Qy 2819 TCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATA 2878
Db 415 CysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIle 434
Qy 2879 AAGACAACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 435 AspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaPro 454
Qy 2908 ---------- 2908
Db 455 CysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 474
Qy 2909 ----------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCT 2956
Db 475 LysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 494
Qy 2957 GTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 495 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhe 514
Qy 3004 ---------- 3004
Db 515 ProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArg 534
Qy 3004 ---------- 3004
Db 535 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 554
Qy 3005 ----------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCT 3043
Db 555 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 574
Qy 3044 TCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATG 3103
Db 575 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 594
Qy 3104 GAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATA 3163
Db 595 GluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 614
Qy 3164 AAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTG 3223
Db 615 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeu 634
Qy 3224 CCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGA 3283
Db 635 ThrLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArg 654
Qy 3284 CCC----------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTT 3328
Db 655 GluGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeu 674
Qy 3329 GAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAG 3388
Db 675 GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln 694
Qy 3389 GTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAG 3448
Db 695 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 714

```
Qy       3449 GAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAA 3508
Db        715 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGlu 734

Qy       3509 AATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACC 3568
Db        735 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 754

Qy       3569 CTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAA 3628
Db        755 LeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 774

Qy       3629 GTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db        775 ValLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 789

RESULT 5
US-09-834-759-475
; Sequence 475, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
;   LENGTH: 1002
;   TYPE: PRT
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(1002)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Alignment Scores:
Pred. No.:              1.74e-182       Length:         1002
Score:                  2367.00         Matches:        511
Percent Similarity:     62.62%          Conservative:   62
Best Local Similarity:  55.85%          Mismatches:     126
Query Match:            36.71%          Indels:         216
DB:                     4               Gaps:           9

US-09-602-362E-26 (1-3673) x US-09-834-759-475 (1-1002)

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258 -
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438
Db         61 GluSerSerThrLysAlaSerAlaAsn---------------------------------- 69

Qy       1439 CTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAAT 1498

Db         69 ---------------------------------------------------------- 69

Qy       1499 CTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCT 1558
Db         69 ---------------------------------------------------------- 69

Qy       1559 ACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCA 1618
Db         70 ----------------------------------------AspGlnArgPheProSer 75

Qy       1619 GAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGT 1678
Db         76 GluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSer 95

Qy       1679 TCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAAT 1738
Db         96 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 115

Qy       1739 AGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAA 1798
Db        116 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 135

Qy       1799 AAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAG 1858
Db        136 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 155

Qy       1859 ATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGT 1918
Db        156 MetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 175

Qy       1919 CCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAA 1978
Db        176 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195

Qy       1979 TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCT 2038
Db        196 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 215

Qy       2039 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2098
Db        216 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 235

Qy       2099 TTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2158
Db        236 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255

Qy       2159 TCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCT 2218
Db        256 SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp------ 273

Qy       2219 GATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCT 2278
Db        273 ---------------------------------------------------------- 273

Qy       2279 TTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAA 2338
Db        274 ----------------------------------GluIleLeuProSerGluSerLys 281

Qy       2339 CAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAG 2398
Db        282 GlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGln 301

Qy       2399 AATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAA 2458
Db        302 LysAspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLys 321

Qy       2459 TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCT 2518
Db        322 LeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 341

Qy       2519 CTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGT 2578
```

```
Db        342 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGlu 361

Qy       2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACA 2638

Db        362 Lys---ProSerAlaPhe-----------GluProAlaIleGluMetGlnLysSerVal 376

Qy       2639 AAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGA 2698

Db        377 ----------------------------------------ProAsnLysAlaLeuGluLeu 383

Qy       2699 CAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATG 2758

Db        384 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro-------------------- 396

Qy       2759 CCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATC 2818

Db        397 ------SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 414

Qy       2819 TCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATA 2878

Db        415 CysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIle 434

Qy       2879 AAGACAACAAATGGCAAAATAGAAGAGTCT---------------------------- 2908

Db        435 AspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaPro 454

Qy       2908 ------------------------------------------------------------ 2908

Db        455 CysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 474

Qy       2909 ------------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCT 2956

Db        475 LysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 494

Qy       2957 GTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA------------ 3004

Db        495 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhe 514

Qy       3004 ------------------------------------------------------------ 3004

Db        515 ProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArg 534

Qy       3004 ------------------------------------------------------------ 3004

Db        535 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 554

Qy       3005 --------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCT 3043

Db        555 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 574

Qy       3044 TCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATG 3103

Db        575 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 594

Qy       3104 GAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATA 3163

Db        595 GluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 614

Qy       3164 AAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTG 3223

Db        615 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeu 634

Qy       3224 CCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGA 3283

Db        635 ThrLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArg 654

Qy       3284 CCC---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTT 3328

Db        655 GluGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeu 674

Qy       3329 GAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAG 3388

Db        675 GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln 694

Qy       3389 GTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAG 3448

Db        695 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 714

Qy       3449 GAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAA 3508

Db        715 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGlu 734

Qy       3509 AATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACC 3568

Db        735 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 754

Qy       3569 CTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAA 3628

Db        755 LeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 774

Qy       3629 GTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673

Db        775 ValLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 789
```

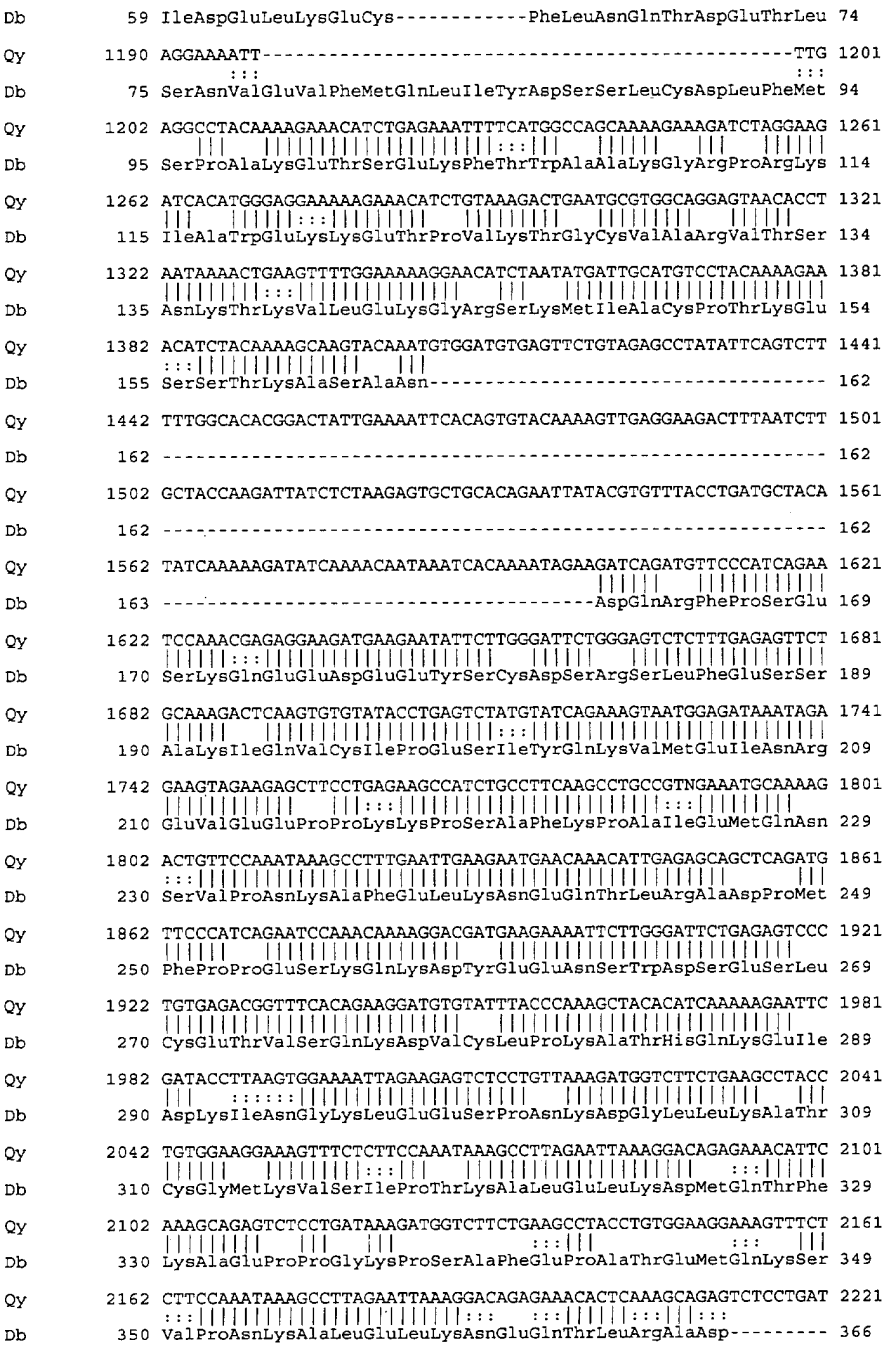
RESULT 6
US-09-620-405B-494
; Sequence 494, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-494

Alignment Scores:

Pred. No.:	8.49e-135	Length:	743
Score:	1775.00	Matches:	414
Percent Similarity:	51.72%	Conservative:	68
Best Local Similarity:	44.42%	Mismatches:	137
Query Match:	27.53%	Indels:	313
DB:	4	Gaps:	13

US-09-602-362E-26 (1-3673) x US-09-620-405B-494 (1-743)

```
Qy        950 CTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACA 1009

Db         24 LeuLeuGluAsnValIleSerLysThrIleAsnProGlnValSerLysThrGlu------ 41

Qy       1010 TCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAA 1069

Db         42 ----------------------------------------------------TyrLys 43

Qy       1070 AGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAA 1129

Db         44 GluLeuLeuGlnGluPheIleAspAspAsnAla--------------ThrThrAsnAla 58

Qy       1130 ATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCT 1189
```


Db 59 IleAspGluLeuLysGluCys----------PheLeuAsnGlnThrAspGluThrLeu 74
Qy 1190 AGGAAAATT----------TTG 1201
Db 75 SerAsnValGluValPheMetGlnLeuIleTyrAspSerSerLeuCysAspLeuPheMet 94
Qy 1202 AGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAG 1261
Db 95 SerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLys 114
Qy 1262 ATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCT 1321
Db 115 IleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSer 134
Qy 1322 AATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAA 1381
Db 135 AsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGlu 154
Qy 1382 ACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTT 1441
Db 155 SerSerThrLysAlaSerAlaAsn---------- 162
Qy 1442 TTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTT 1501
Db 162 ---------- 162
Qy 1502 GCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACA 1561
Db 162 ---------- 162
Qy 1562 TATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAA 1621
Db 163 ----------AspGlnArgPheProSerGlu 169
Qy 1622 TCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCT 1681
Db 170 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSer 189
Qy 1682 GCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGA 1741
Db 190 AlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArg 209
Qy 1742 GAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAG 1801
Db 210 GluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsn 229
Qy 1802 ACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATG 1861
Db 230 SerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMet 249
Qy 1862 TTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCC 1921
Db 250 PheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLeu 269
Qy 1922 TGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTC 1981
Db 270 CysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIle 289
Qy 1982 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACC 2041
Db 290 AspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThr 309
Qy 2042 TGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTC 2101
Db 310 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329
Qy 2102 AAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT 2161
Db 330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349
Qy 2162 CTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGAT 2221
Db 350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp--------- 366
Qy 2222 AATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTA 2281
Db 366 ---------- 366
Qy 2282 GAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAA 2341
Db 367 ----------GluIleLeuProSerGluSerLysGln 375
Qy 2342 AAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAAT 2401
Db 376 LysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGlnLys 395
Qy 2402 GATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTA 2461
Db 396 AspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLysLeu 415
Qy 2462 GAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTT 2521
Db 416 GluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIle 435
Qy 2522 CCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCT 2581
Db 436 ProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys 455
Qy 2582 GTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAA 2641
Db 456 ---ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal--- 469
Qy 2642 GTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAA 2701
Db 470 ----------ProAsnLysAlaLeuGluLeuLys 477
Qy 2702 CAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCC 2761
Db 478 AsnGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 489
Qy 2762 ACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCT 2821
Db 490 ---SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 508
Qy 2822 GTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAG 2881
Db 509 GluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAsp 528
Qy 2882 ACAACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 529 LysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCys 548
Qy 2908 ---------- 2908
Db 549 ArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 568
Qy 2909 ---------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTT 2959
Db 569 AlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 588
Qy 2960 CCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 589 ProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePro 608
Qy 3004 ---------- 3004
Db 609 SerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArgGlu 628
Qy 3004 ---------- 3004
Db 629 ThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLys 648
Qy 3005 ----------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCT 3046
Db 649 IleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSer 668

```
Qy       3047 TGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAA 3106
Db        669 CysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGlu 688

Qy       3107 CAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA 3166
Db        689 GlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLys 708

Qy       3167 TCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCT 3226
Db        709 SerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal--------- 725

Qy       3227 TTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC 3286
Db        725 ------------------------------------------------------------ 725

Qy       3287 GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATA 3346
Db        725 ------------------------------------------------------------ 725

Qy       3347 CAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAA 3406
Db        726 -------------------------------------Arg-PheLeuThrLeuMetLy 732

Qy       3407 AGTGAAAATGATCTCTTTCATGAAAATTGCATGT 3440
Db        732 sMetLysIleIleSerTyrMetLysIleAlaCys 743
```

RESULT 7
US-09-834-759-494
; Sequence 494, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-494

Alignment Scores:

Pred. No.:	8.49e-135	Length:	743
Score:	1775.00	Matches:	414
Percent Similarity:	51.72%	Conservative:	68
Best Local Similarity:	44.42%	Mismatches:	137
Query Match:	27.53%	Indels:	313
DB:	4	Gaps:	13

US-09-602-362E-26 (1-3673) x US-09-834-759-494 (1-743)

```
Qy        950 CTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACA 1009
Db         24 LeuLeuGluAsnValIleSerLysThrIleAsnProGlnValSerLysThrGlu------ 41

Qy       1010 TCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAA 1069

Db         42 ----------------------------------------------------TyrLys 43

Qy       1070 AGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAA 1129
Db         44 GluLeuLeuGlnGluPheIleAspAspAsnAla--------------ThrThrAsnAla 58

Qy       1130 ATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCT 1189
Db         59 IleAspGluLeuLysGluCys------------PheLeuAsnGlnThrAspGluThrLeu 74

Qy       1190 AGGAAAAATT----------------------------------------------TTG 1201
Db         75 SerAsnValGluValPheMetGlnLeuIleTyrAspSerSerLeuCysAspLeuPheMet 94

Qy       1202 AGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAG 1261
Db         95 SerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLys 114

Qy       1262 ATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCT 1321
Db        115 IleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSer 134

Qy       1322 AATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAA 1381
Db        135 AsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGlu 154

Qy       1382 ACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTT 1441
Db        155 SerSerThrLysAlaSerAlaAsn---------------------------------- 162

Qy       1442 TTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTT 1501
Db        162 ------------------------------------------------------------ 162

Qy       1502 GCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACA 1561
Db        162 ------------------------------------------------------------ 162

Qy       1562 TATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAA 1621
Db        163 -------------------------------------AspGlnArgPheProSerGlu 169

Qy       1622 TCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCT 1681
Db        170 SerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSer 189

Qy       1682 GCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGA 1741
Db        190 AlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArg 209

Qy       1742 GAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAG 1801
Db        210 GluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsn 229

Qy       1802 ACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATG 1861
Db        230 SerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMet 249

Qy       1862 TTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCC 1921
Db        250 PheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLeu 269

Qy       1922 TGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTC 1981
Db        270 CysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIle 289

Qy       1982 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACC 2041
Db        290 AspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThr 309

Qy       2042 TGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTC 2101
Db        310 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329
```

```
Qy       2102 AAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT 2161
Db        330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349

Qy       2162 CTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGAT 2221
Db        350 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp--------- 366

Qy       2222 AATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTA 2281
Db        366 ------------------------------------------------------------ 366

Qy       2282 GAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAA 2341
Db        367 --------------------------------GluIleLeuProSerGluSerLysGln 375

Qy       2342 AAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAAT 2401
Db        376 LysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGlnLys 395

Qy       2402 GATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTA 2461
Db        396 AspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLysLeu 415

Qy       2462 GAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTT 2521
Db        416 GluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIle 435

Qy       2522 CCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCT 2581
Db        436 ProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys 455

Qy       2582 GTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAA 2641
Db        456 ---ProSerAlaPhe-----------GluProAlaIleGluMetGlnLysSerVal--- 469

Qy       2642 GTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAA 2701
Db        470 ----------------------------------ProAsnLysAlaLeuGluLeuLys 477

Qy       2702 CAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCC 2761
Db        478 AsnGluGlnThrLeuArgAlaAspGluIleLeuPro----------------------- 489

Qy       2762 ACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCT 2821
Db        490 ---SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 508

Qy       2822 GTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAG 2881
Db        509 GluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIleAsp 528

Qy       2882 ACAACAAATGGCAAAATAGAAGAGTCT---------------------------------- 2908
Db        529 LysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCys 548

Qy       2908 ------------------------------------------------------------ 2908
Db        549 ArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 568

Qy       2909 ---------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTT 2959
Db        569 AlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 588

Qy       2960 CCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA--------------- 3004
Db        589 ProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePro 608

Qy       3004 ------------------------------------------------------------ 3004
Db        609 SerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArgGlu 628

Qy       3004 ------------------------------------------------------------ 3004
Db        629 ThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLys 648

Qy       3005 -----------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCT 3046
Db        649 IleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSer 668

Qy       3047 TGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAA 3106
Db        669 CysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGlu 688

Qy       3107 CAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA 3166
Db        689 GlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLys 708

Qy       3167 TCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCT 3226
Db        709 SerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal--------- 725

Qy       3227 TTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC 3286
Db        725 ------------------------------------------------------------ 725

Qy       3287 GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATA 3346
Db        725 ------------------------------------------------------------ 725

Qy       3347 CAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAA 3406
Db        726 ------------------------------------Arg-PheLeuThrLeuMetLy 732

Qy       3407 AGTGAAAATGATCTCTTTCATGAAAATTGCATGT 3440
Db        732 sMetLysIleIleSerTyrMetLysIleAlaCys 743
```

RESULT 8
US-09-620-405B-469
; Sequence 469, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-469

Alignment Scores:

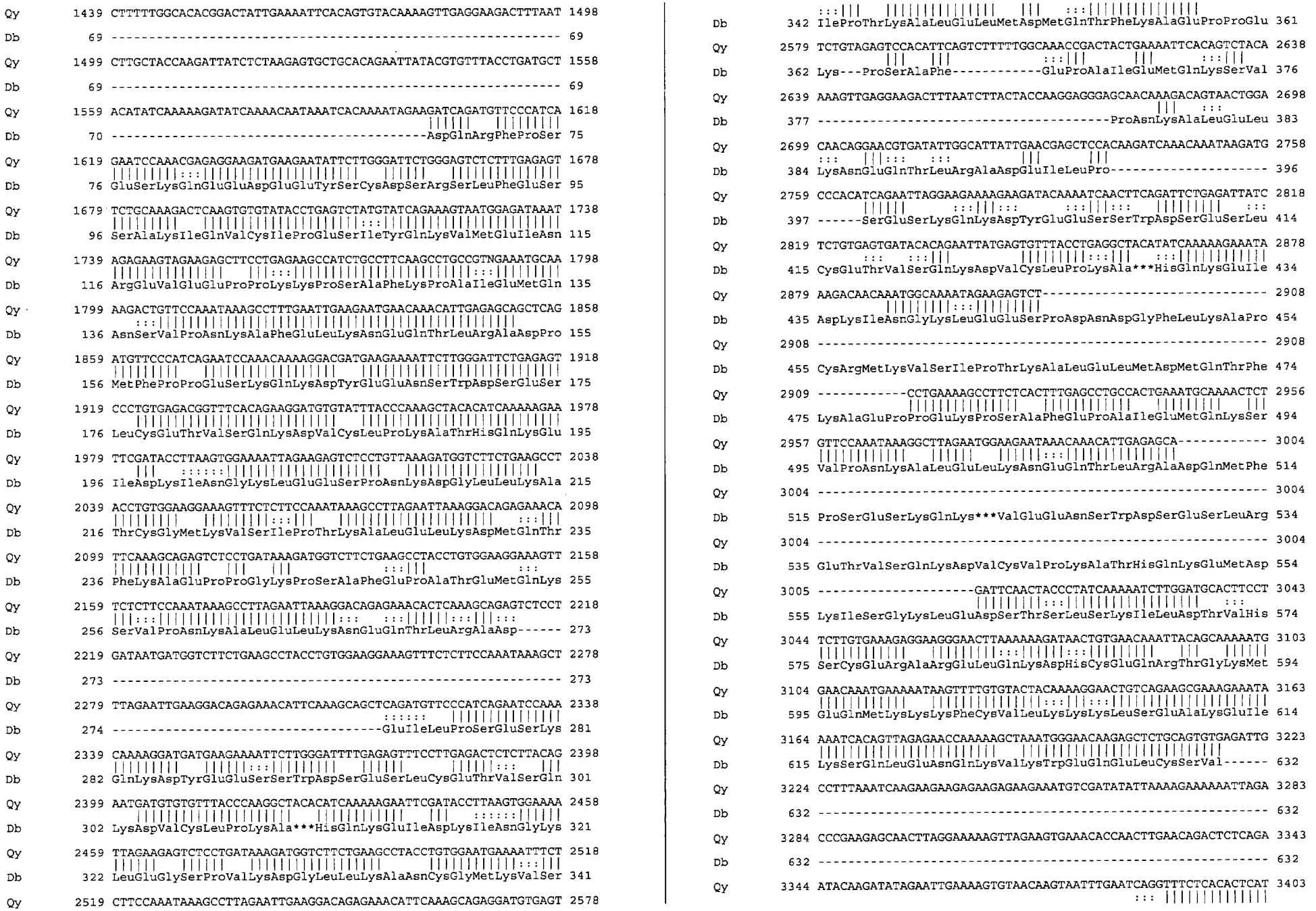
Pred. No.:	2.24e-134	Length:	650
Score:	1769.50	Matches:	396
Percent Similarity:	54.38%	Conservative:	57
Best Local Similarity:	47.54%	Mismatches:	112
Query Match:	27.44%	Indels:	268
DB:	4	Gaps:	9

US-09-602-362E-26 (1-3673) x US-09-620-405B-469 (1-650)

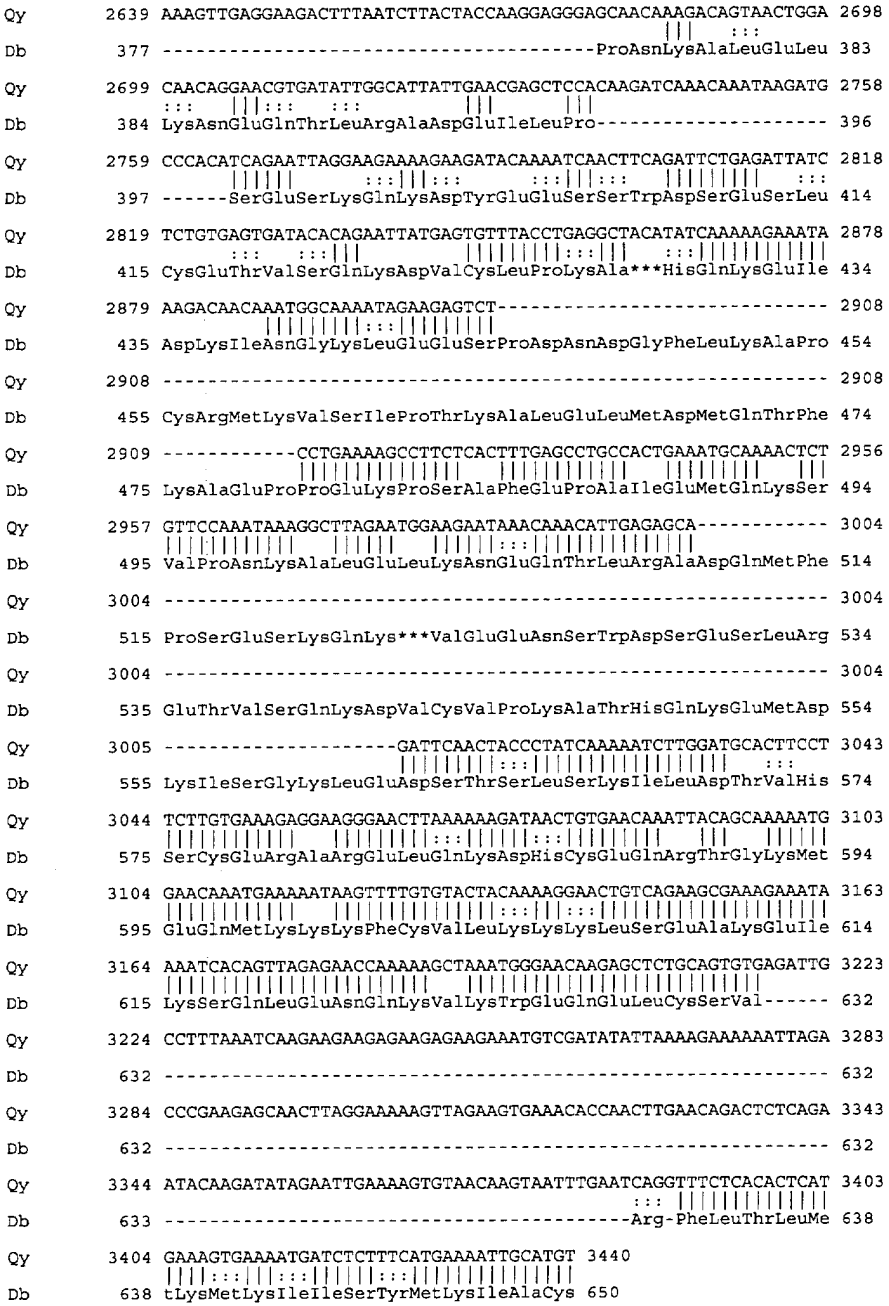
```
Db        595 GluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 614

Qy       3164 AAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTG 3223

Db        615 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal------ 632

Qy       3224 CCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGA 3283

Db        632 ---------------------------------------------------------- 632

Qy       3284 CCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGA 3343

Db        632 ---------------------------------------------------------- 632

Qy       3344 ATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCAT 3403

Db        633 ----------------------------------------Arg-PheLeuThrLeuMe 638

Qy       3404 GAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGT 3440

Db        638 tLysMetLysIleIleSerTyrMetLysIleAlaCys 650

RESULT 9
US-09-433-826B-469
; Sequence 469, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION:  COMPOSITIONS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.470C4
;  CURRENT APPLICATION NUMBER: US/09/433,826B
;  CURRENT FILING DATE:  1999-11-03
;  NUMBER OF SEQ ID NOS:  474
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 469
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (310)
;   NAME/KEY: Xaa = Any Amino Acid<221> unsure
;   LOCATION: (429)
;   NAME/KEY: Xaa = Any Amino Acid<221> unsure
;   LOCATION: (522)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-433-826B-469

Alignment Scores:
Pred. No.:              2.24e-134       Length:         650
Score:                  1769.50         Matches:        396
Percent Similarity:     54.38%          Conservative:   57
Best Local Similarity:  47.54%          Mismatches:     112
Query Match:            27.44%          Indels:         268
DB:                     4               Gaps:           9

US-09-602-362E-26 (1-3673) x US-09-433-826B-469 (1-650)

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258

Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318

Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378

Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438

Db         61 GluSerSerThrLysAlaSerAlaAsn------------------------------- 69

Qy       1439 CTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAAT 1498

Db         69 ---------------------------------------------------------- 69

Qy       1499 CTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCT 1558

Db         69 ---------------------------------------------------------- 69

Qy       1559 ACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCA 1618

Db         70 ----------------------------------------AspGlnArgPheProSer 75

Qy       1619 GAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGT 1678

Db         76 GluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSer 95

Qy       1679 TCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAAT 1738

Db         96 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 115

Qy       1739 AGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAA 1798

Db        116 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 135

Qy       1799 AAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAG 1858

Db        136 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 155

Qy       1859 ATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGT 1918

Db        156 MetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 175

Qy       1919 CCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAA 1978

Db        176 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195

Qy       1979 TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCT 2038

Db        196 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 215

Qy       2039 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2098

Db        216 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 235

Qy       2099 TTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2158

Db        236 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255

Qy       2159 TCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCT 2218

Db        256 SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp------ 273

Qy       2219 GATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCT 2278

Db        273 ---------------------------------------------------------- 273

Qy       2279 TTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAA 2338

Db        274 ----------------------------------GluIleLeuProSerGluSerLys 281

Qy       2339 CAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAG 2398

Db        282 GlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGln 301

Qy       2399 AATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAA 2458
```

```
Db        302 LysAspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLys 321

Qy       2459 TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCT 2518
              ||||||  ||||||  ||||||||||||||||||  ||||||||||||:::|||
Db        322 LeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 341

Qy       2519 CTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGT 2578
              :::|||  ||||||||||||||  |||  :::||||||||||||||||
Db        342 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGlu 361

Qy       2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACA 2638
              |||  |||           :::|||      |||     :::|||
Db        362 Lys---ProSerAlaPhe-----------GluProAlaIleGluMetGlnLysSerVal 376

Qy       2639 AAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGA 2698
                                                        |||  :::
Db        377 ------------------------------------ProAsnLysAlaLeuGluLeu 383

Qy       2699 CAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATG 2758
              :::  |||:::  :::        |||      |||
Db        384 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro------------------- 396

Qy       2759 CCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATC 2818
              ||||||      :::|||:::      :::|||:::   |||||||||  :::
Db        397 ------SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 414

Qy       2819 TCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATA 2878
              :::   :::|||         |||||||||:::|||   :::||||||||||||
Db        415 CysGluThrValSerGlnLysAspValCysLeuProLysAla***HisGlnLysGluIle 434

Qy       2879 AAGACAACAAATGGCAAAATAGAAGAGTCT---------------------------- 2908
                    |||||||||:::|||||||||
Db        435 AspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaPro 454

Qy       2908 ------------------------------------------------------------ 2908

Db        455 CysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 474

Qy       2909 ------------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCT 2956
                          ||||||||||||||   ||||||||||||   |||||||||   |||
Db        475 LysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 494

Qy       2957 GTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA------------ 3004
              ||||||||||   ||||||   ||||||:::||||||||||||||
Db        495 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhe 514

Qy       3004 ------------------------------------------------------------ 3004

Db        515 ProSerGluSerLysGlnLys***ValGluGluAsnSerTrpAspSerGluSerLeuArg 534

Qy       3004 ------------------------------------------------------------ 3004

Db        535 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 554

Qy       3005 --------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCT 3043
                                 |||||||||:::|||||||||||||||||   :::
Db        555 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 574

Qy       3044 TCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATG 3103
              ||||||||||   |||||||||:::||||||:::|||||||||   |||   ||||||
Db        575 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 594

Qy       3104 GAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATA 3163
              ||||||||||   ||||||||||||||:::|||:::||||||||||||||||||||||
Db        595 GluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 614

Qy       3164 AAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTG 3223
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        615 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal------ 632

Qy       3224 CCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGA 3283

Db        632 ------------------------------------------------------------ 632
```

```
Qy       3284 CCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGA 3343

Db        632 ------------------------------------------------------------ 632

Qy       3344 ATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCAT 3403
                                                                :::  ||||||||||||||
Db        633 ----------------------------------------Arg-PheLeuThrLeuMe 638

Qy       3404 GAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGT 3440
              ||||:::|||:::||||||:::||||||||||||||||
Db        638 tLysMetLysIleIleSerTyrMetLysIleAlaCys 650

RESULT 10
US-09-604-287A-469
; Sequence 469, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C7
;  CURRENT APPLICATION NUMBER: US/09/604,287A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (310)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (429)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (522)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-469

Alignment Scores:
Pred. No.:              2.24e-134       Length:         650
Score:                  1769.50         Matches:        396
Percent Similarity:     54.38%          Conservative:   57
Best Local Similarity:  47.54%          Mismatches:     112
Query Match:            27.44%          Indels:         268
DB:                     4               Gaps:           9

US-09-602-362E-26 (1-3673) x US-09-604-287A-469 (1-650)

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
              :::  |||  |||||||||||||||||||:::|||  ||||||  |||  |||
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
              ||||||  ||||||:::|||||||||  |||||||||  |||||||||  ||||||
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
                 |||||||||:::|||||||||||||||  |||  ||||||||||||||||||||
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438
              |||:::|||||||||||||||  |||
Db         61 GluSerSerThrLysAlaSerAlaAsn-------------------------------- 69
```



```
Db        633 ----------------------------------------Arg-PheLeuThrLeuMe 638

Qy       3404 GAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGT 3440
              ||||:::|||:::||||||:::|||||||||||||||
Db        638 tLysMetLysIleIleSerTyrMetLysIleAlaCys 650

RESULT 11
US-09-834-759-469
; Sequence 469, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 469
;   LENGTH: 650
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (310)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (429)
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   NAME/KEY: unsure
;   LOCATION: (522)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-469

Alignment Scores:
Pred. No.:              2.24e-134       Length:         650
Score:                  1769.50         Matches:        396
Percent Similarity:     54.38%          Conservative:   57
Best Local Similarity:  47.54%          Mismatches:     112
Query Match:            27.44%          Indels:         268
DB:                     4               Gaps:           9

US-09-602-362E-26 (1-3673) x US-09-834-759-469 (1-650)

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
              :::   |||   |||||||||||||||||||||:::|||   ||||||   |||   |||
Db          1 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 20

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
              ||||||   ||||||:::||||||||   ||||||||   ||||||||   ||||||
Db         21 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
              |||||||||:::|||||||||||||||   |||   ||||||||||||||||||||
Db         41 SerAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLys 60

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438
              |||:::|||||||||||||||   |||
Db         61 GluSerSerThrLysAlaSerAlaAsn------------------------------ 69

Qy       1439 CTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAAT 1498

Db         69 ------------------------------------------------------------ 69

Qy       1499 CTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCT 1558

Db         69 ------------------------------------------------------------ 69

Qy       1559 ACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCA 1618
                                                          ||||||   |||||||||
Db         70 ----------------------------------------AspGlnArgPheProSer 75

Qy       1619 GAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGT 1678
              |||||||||:::|||||||||||||||||||||   ||||||   |||||||||||||||
Db         76 GluSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSer 95

Qy       1679 TCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAAT 1738
              |||||||||   |||||||||||||||||||||:::|||||||||||||||||||||||
Db         96 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 115

Qy       1739 AGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAA 1798
              |||||||||||||||   |||:::|||||||||||||||||||||:::|||||||||
Db        116 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 135

Qy       1799 AAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAG 1858
                 :::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 155

Qy       1859 ATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGT 1918
              |||||||||   |||||||||||||||||   |||||||||||||||||||||||||||||
Db        156 MetPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 175

Qy       1919 CCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAA 1978
              ||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db        176 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195

Qy       1979 TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCT 2038
                 |||   ::::::|||||||||||||||||||||   ||||||||||||||||||
Db        196 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 215

Qy       2039 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2098
              |||||||||   |||||||||:::|||   |||||||||||||||||||||   :::|||
Db        216 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 235

Qy       2099 TTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2158
              ||||||||||||   |||   |||              :::|||              :::
Db        236 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255

Qy       2159 TCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCT 2218
              |||:::|||||||||||||||||||||:::   :::||||||:::|||:::
Db        256 SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp------ 273

Qy       2219 GATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCT 2278

Db        273 ------------------------------------------------------------ 273

Qy       2279 TTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAA 2338
                                                      ::::::   ||||||||||||||
Db        274 ----------------------------------GluIleLeuProSerGluSerLys 281

Qy       2339 CAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAG 2398
              |||||||||   ||||||:::|||||||||   ||||||       ||||||:::   |||
Db        282 GlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGln 301

Qy       2399 AATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAA 2458
              |||||||||||||||||||   ||||||||||||   |||   ::::::||||||
Db        302 LysAspValCysLeuProLysAla***HisGlnLysGluIleAspLysIleAsnGlyLys 321

Qy       2459 TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCT 2518
              ||||||   ||||||   ||||||||||||||||||       ||||||||||:::|||
Db        322 LeuGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 341

Qy       2519 CTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGT 2578
              :::|||   |||||||||||||||   |||   :::|||||||||||||||
Db        342 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGlu 361

Qy       2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACA 2638
                       |||   |||              :::|||       |||      :::|||
Db        362 Lys---ProSerAlaPhe------------GluProAlaIleGluMetGlnLysSerVal 376
```

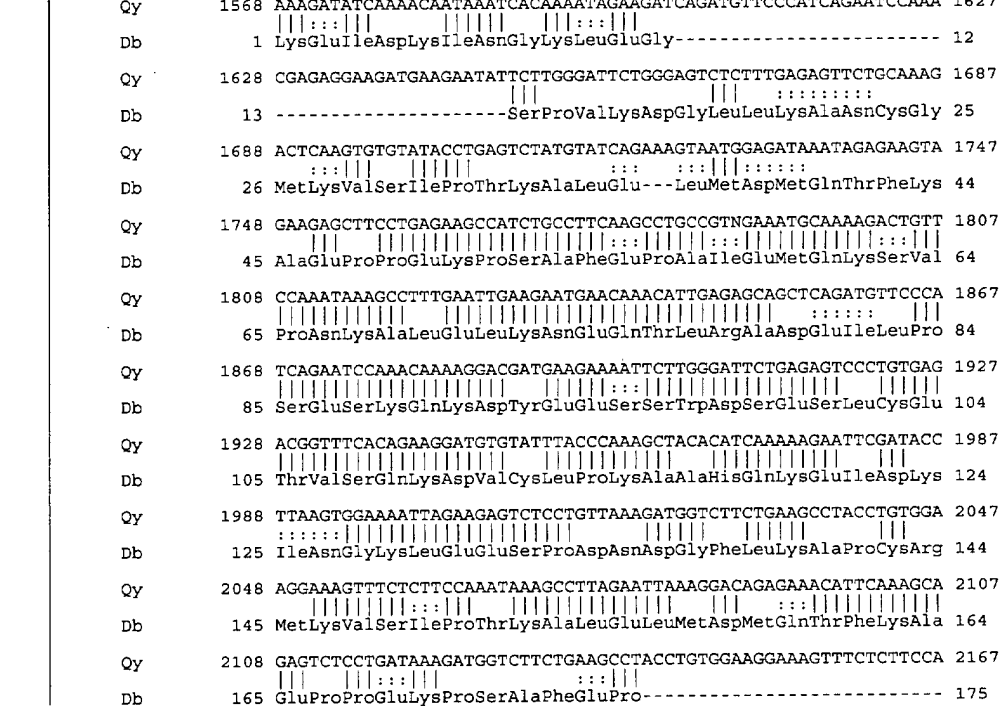


RESULT 12
US-09-620-405B-473
; Sequence 473, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-620-405B-473

Alignment Scores:

Pred. No.:	8.19e-106	Length:	445
Score:	1414.50	Matches:	315
Percent Similarity:	53.27%	Conservative:	43
Best Local Similarity:	46.88%	Mismatches:	79
Query Match:	21.94%	Indels:	235
DB:	4	Gaps:	6

US-09-602-362E-26 (1-3673) x US-09-620-405B-473 (1-445)



```
Qy       2168 AATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAATGAT 2227
                 |||:::|||::::::
Db        176 ------AlaIleGluMetGlnLys---------------------------------- 181

Qy       2228 GGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTG 2287
                                              |||:::|||||||||||||||||||||
Db        182 --------------------------------SerValProAsnLysAlaLeuGluLeu 190

Qy       2288 AAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGAT 2347
                 |||:::   :::|||   :::|||   ||||||||||||||||||||||||||||
Db        191 LysAsnGluGlnThrLeuArgAlaAspGlnMetProSerGluSerLysGlnLysLys 210

Qy       2348 GATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGATGTG 2407
                 ||||||||||||||||||   ||||||      ||||||:::   |||   ||||||
Db        211 ValGluGluAsnSerTrpAspSerGluSerLeuArgGluThrValSerGlnLysAspVal 230

Qy       2408 TGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG 2467
                 |||:::|||||||||||||||||||||||||   |||   :::|||||||||||||||
Db        231 CysValProLysAlaThrHisGlnLysGluMetAspLysIleSerGlyLysLeuGlu--- 249

Qy       2468 TCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAAT 2527
Db        249 ------------------------------------------------------------ 249

Qy       2528 AAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAG 2587
Db        249 ------------------------------------------------------------ 249

Qy       2588 TCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAG 2647
Db        249 ------------------------------------------------------------ 249

Qy       2648 GAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAA 2707
Db        249 ------------------------------------------------------------ 249

Qy       2708 CGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA 2767
Db        249 ------------------------------------------------------------ 249

Qy       2768 GAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGT 2827
Db        249 ------------------------------------------------------------ 249

Qy       2828 GATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACA 2887
Db        249 ------------------------------------------------------------ 249

Qy       2888 AATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATG 2947
Db        249 ------------------------------------------------------------ 249

Qy       2948 CAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGAT 3007
                                                                        |||
Db        250 ------------------------------------------------------------Asp 250

Qy       3008 TCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTT 3067
                 ||||||:::|||||||||||||||||||   :::   ||||||||||||   |||||||||
Db        251 SerThrSerLeuSerLysIleLeuAspThrValHisSerCysGluArgAlaArgGluLeu 270

Qy       3068 AAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGT 3127
                 :::||||||:::|||||||||   |||   ||||||||||||||||||   |||||||||
Db        271 GlnLysAspHisCysGluGlnArgThrGlyLysMetGluGlnMetLysLysLysPheCys 290

Qy       3128 GTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAA 3187
                 ||||||:::|||:::|||||||||||||||||||||||||||||||||||||||||||
Db        291 ValLeuLysLysLysLeuSerGluAlaLysGluIleLysSerGlnLeuGluAsnGlnLys 310

Qy       3188 GCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAG 3247
                 ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        311 ValLysTrpGluGlnGluLeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLys 330

Qy       3248 AGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC----------GAAGAG 3292
                 |||||||||   |||||||||   ||||||||||||                ||||||
Db        331 ArgArgAsnAlaAspIleLeuAsnGluLysIleArgGluGluLeuGlyArgIleGluGlu 350

Qy       3293 CAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGAT 3352
                 |||   ||||||:::|||||||||||   ||||||||||||   ||||||||||||||
Db        351 GlnHisArgLysGluLeuGluValLysGlnGlnLeuGluGlnAlaLeuArgIleGlnAsp 370

Qy       3353 ATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAA 3412
                 ||||||||||||||||||   ||||||||||||||||||||||||||||||:::|||
Db        371 IleGluLeuLysSerValGluSerAsnLeuAsnGlnValSerHisThrHisGluAsnGlu 390

Qy       3413 AATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAA 3472
                 |||   |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGlu 410

Qy       3473 GTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAG 3532
                 :::|||||||||||||||||:::|||   ||||||||||||||||||||||||||||
Db        411 IleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLys 430

Qy       3533 ATTTTACAAGAAAAGAATGCTGAACTTCAAATGACC 3568
                 ||||||:::|||||||||||||||||||||||||
Db        431 IleLeuLysGluLysAsnAlaGluLeuGlnMetThr 442
```

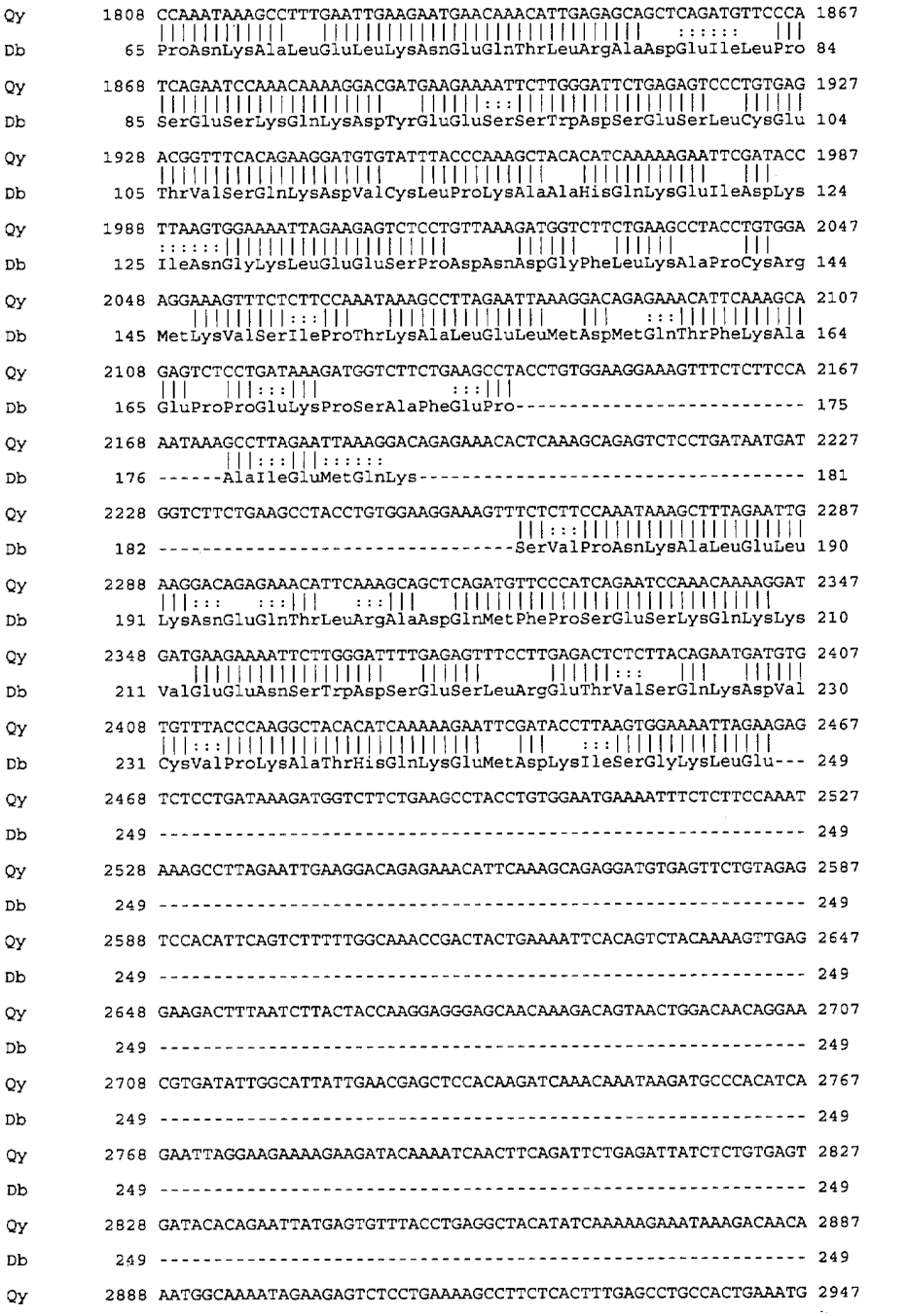
RESULT 13
US-09-433-826B-473
; Sequence 473, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-826B-473

Alignment Scores:

Pred. No.:	8.19e-106	Length:	445
Score:	1414.50	Matches:	315
Percent Similarity:	53.27%	Conservative:	43
Best Local Similarity:	46.88%	Mismatches:	79
Query Match:	21.94%	Indels:	235
DB:	4	Gaps:	6

US-09-602-362E-26 (1-3673) x US-09-433-826B-473 (1-445)

```
Qy       1568 AAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAATCCAAA 1627
                 |||:::|||      ||||||   |||:::|||
Db          1 LysGluIleAspLysIleAsnGlyLysLeuGluGly------------------------ 12

Qy       1628 CGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAG 1687
                                          |||      |||   :::::::::
Db         13 ---------------------SerProValLysAspGlyLeuLeuLysAlaAsnCysGly 25

Qy       1688 ACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGAAGTA 1747
                    :::|||   ||||||          :::   :::|||::::::
Db         26 MetLysValSerIleProThrLysAlaLeuGlu---LeuMetAspMetGlnThrPheLys 44

Qy       1748 GAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGACTGTT 1807
                 |||   |||||||||||||||||:::||||||:::||||||||||||:::|||
Db         45 AlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 64
```

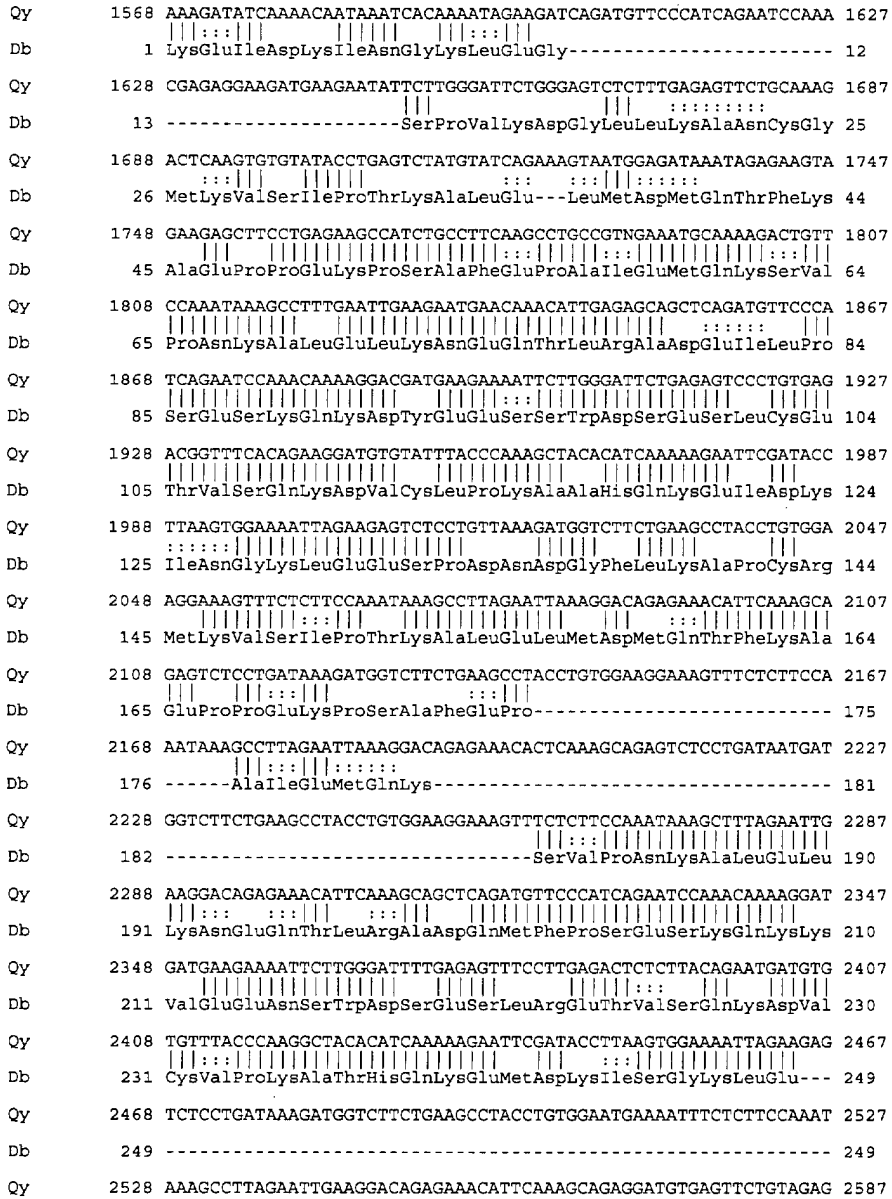


RESULT 14
US-09-604-287A-473
; Sequence 473, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-287A-473

Alignment Scores:
Pred. No.: 8.19e-106 Length: 445

Score: 1414.50 Matches: 315
Percent Similarity: 53.27% Conservative: 43
Best Local Similarity: 46.88% Mismatches: 79
Query Match: 21.94% Indels: 235
DB: 4 Gaps: 6

US-09-602-362E-26 (1-3673) x US-09-604-287A-473 (1-445)



RESULT 15
US-09-834-759-473

```
; Sequence 473, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C9
;  CURRENT APPLICATION NUMBER: US/09/834,759
;  CURRENT FILING DATE:  2001-04-13
;  NUMBER OF SEQ ID NOS: 547
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 473
;   LENGTH: 445
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-759-473

Alignment Scores:
Pred. No.:                8.19e-106      Length:          445
Score:                    1414.50        Matches:         315
Percent Similarity:       53.27%         Conservative:    43
Best Local Similarity:    46.88%         Mismatches:      79
Query Match:              21.94%         Indels:          235
DB:                       4              Gaps:            6

US-09-602-362E-26 (1-3673) x US-09-834-759-473 (1-445)

Qy       1568 AAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAATCCAAA 1627
                 |||:::|||      ||||||   |||:::|||
Db          1 LysGluIleAspLysIleAsnGlyLysLeuGluGly------------------------ 12

Qy       1628 CGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAG 1687
                              |||              |||   :::::::::
Db         13 ---------------------SerProValLysAspGlyLeuLeuLysAlaAsnCysGly 25

Qy       1688 ACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGAAGTA 1747
                 :::|||   ||||||            :::   :::|||::::::
Db         26 MetLysValSerIleProThrLysAlaLeuGlu---LeuMetAspMetGlnThrPheLys 44

Qy       1748 GAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGACTGTT 1807
                 |||   |||||||||||||||||||||:::||||||:::|||||||||||||:::|||
Db         45 AlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 64

Qy       1808 CCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATGTTCCCA 1867
                 ||||||||||||   |||||||||||||||||||||||||||||   ::::::   |||
Db         65 ProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro 84

Qy       1868 TCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCCTGTGAG 1927
                 |||||||||||||||||||   ||||||:::|||||||||||||||||   ||||||
Db         85 SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGlu 104

Qy       1928 ACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGATACC 1987
                 |||||||||||||||||||   ||||||||||||   ||||||||||||   |||
Db        105 ThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGlnLysGluIleAspLys 124

Qy       1988 TTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGTGGA 2047
                 ::::::|||||||||||||||||||||||      ||||||   ||||||      |||
Db        125 IleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCysArg 144

Qy       2048 AGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAAAGCA 2107
                 |||||||||:::|||   ||||||||||||||   |||   :::|||||||||||||
Db        145 MetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAla 164

Qy       2108 GAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCA 2167
                 |||   |||:::|||              :::|||
Db        165 GluProProGluLysProSerAlaPheGluPro-------------------------- 175

Qy       2168 AATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAATGAT 2227
                      |||:::|||::::::
Db        176 ------AlaIleGluMetGlnLys---------------------------------- 181

Qy       2228 GGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTG 2287
                                                  |||:::|||||||||||||||||||||
Db        182 --------------------------------SerValProAsnLysAlaLeuGluLeu 190

Qy       2288 AAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGAT 2347
                 |||:::   :::|||   :::|||   |||||||||||||||||||||||||||||||
Db        191 LysAsnGluGlnThrLeuArgAlaAspGlnMetPheProSerGluSerLysGlnLysLys 210

Qy       2348 GATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGATGTG 2407
                    |||||||||||||||||   ||||||   ||||||:::   |||   ||||||
Db        211 ValGluGluAsnSerTrpAspSerGluSerLeuArgGluThrValSerGlnLysAspVal 230

Qy       2408 TGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG 2467
                 |||:::|||||||||||||||||||||||||   |||   :::|||||||||||||||
Db        231 CysValProLysAlaThrHisGlnLysGluMetAspLysIleSerGlyLysLeuGlu--- 249

Qy       2468 TCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAAT 2527
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2528 AAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAG 2587
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2588 TCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAG 2647
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2648 GAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAA 2707
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2708 CGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA 2767
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2768 GAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGT 2827
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2828 GATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACA 2887
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2888 AATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATG 2947
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 ------------------------------------------------------------ 249

Qy       2948 CAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGAT 3007
                                                                        |||
Db        250 ------------------------------------------------------------Asp 250

Qy       3008 TCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTT 3067
                 ||||||:::|||||||||||||||||||   :::   ||||||||||||   |||||||||
Db        251 SerThrSerLeuSerLysIleLeuAspThrValHisSerCysGluArgAlaArgGluLeu 270

Qy       3068 AAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGT 3127
                 :::||||||:::|||||||||   |||   |||||||||||||||||   |||||||||
Db        271 GlnLysAspHisCysGluGlnArgThrGlyLysMetGluGlnMetLysLysLysPheCys 290

Qy       3128 GTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAA 3187
                 ||||||:::|||:::||||||||||||||||||||||||||||||||||||||||||||
Db        291 ValLeuLysLysLysLeuSerGluAlaLysGluIleLysSerGlnLeuGluAsnGlnLys 310

Qy       3188 GCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAG 3247
                 ||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        311 ValLysTrpGluGlnGluLeuCysSerValArgLeuThrLeuAsnGlnGluGluGluLys 330

Qy       3248 AGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC--------------GAAGAG 3292
```

```
                ||||||||||  |||||||||  ||||||||||||                  ||||||
Db        331 ArgArgAsnAlaAspIleLeuAsnGluLysIleArgGluGluLeuGlyArgIleGluGlu 350

Qy       3293 CAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGAT 3352
                |||    ||||||:::||||||||||||    ||||||||||||   ||||||||||||||||
Db        351 GlnHisArgLysGluLeuGluValLysGlnGlnLeuGluGlnAlaLeuArgIleGlnAsp 370

Qy       3353 ATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAA 3412
                ||||||||||||||||    ||||||||||||||||||||||||||||||:::|||
Db        371 IleGluLeuLysSerValGluSerAsnLeuAsnGlnValSerHisThrHisGluAsnGlu 390

Qy       3413 AATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAA 3472
                |||    |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGlu 410

Qy       3473 GTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAG 3532
                :::||||||||||||||||||:::|||   ||||||||||||||||||||||||||||
Db        411 IleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLys 430

Qy       3533 ATTTTACAAGAAAAGAATGCTGAACTTCAAATGACC 3568
                ||||||:::|||||||||||||||||||||||||||
Db        431 IleLeuLysGluLysAsnAlaGluLeuGlnMetThr 442


Search completed: July 15, 2004, 09:29:05
Job time : 94.2381 secs
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 08:46:47 ; Search time 172.464 Seconds
(without alignments)
13439.296 Million cell updates/sec

Title: US-09-602-362E-26
Perfect score: 6448
Sequence: 1 caagagcttggcgatacaga.....tgacttctaaattgaaggaa 3673

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093626_22001/app_query.fasta_1.10325
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09602362_@CGN_1_1_814_@runat_15072004_093626_22001 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5173	80.2	1011	4	Q9BXX2	Q9bxx2 homo sapien
2	3458	53.6	1341	4	Q9BXX3	Q9bxx3 homo sapien
3	1257	19.5	1709	4	Q9UPS8	Q9ups8 homo sapien
4	1250.5	19.4	1710	4	Q9H1Q1	Q9h1q1 homo sapien
5	1011.5	15.7	424	4	Q9NSI9	Q9nsi9 homo sapien
6	701	10.9	823	4	Q9H0H6	Q9h0h6 homo sapien
7	606	9.4	996	4	Q8IVF6	Q8ivf6 homo sapien
8	598.5	9.3	546	4	Q9NW69	Q9nw69 homo sapien
9	566.5	8.8	733	11	Q811D2	Q811d2 mus musculu
10	553	8.6	168	4	Q8N7B8	Q8n7b8 homo sapien
11	509	7.9	533	11	Q9CS61	Q9cs61 mus musculu
12	508.5	7.9	264	4	Q9H560	Q9h560 homo sapien
13	499	7.7	357	4	Q7Z468	Q7z468 homo sapien
14	489	7.6	718	4	Q9HCD1	Q9hcd1 homo sapien
15	487	7.6	279	11	Q9D504	Q9d504 mus musculu
16	479	7.4	279	11	Q810N2	Q810n2 mus musculu
17	475.5	7.4	280	11	Q8BVP4	Q8bvp4 mus musculu
18	469	7.3	239	4	Q8N2N9	Q8n2n9 homo sapien
19	440	6.8	435	11	Q9CUQ9	Q9cuq9 mus musculu
20	440	6.8	1413	11	Q8CGB3	Q8cgb3 mus musculu
21	439.5	6.8	1416	4	Q9HCL1	Q9hcl1 homo sapien
22	438.5	6.8	347	11	Q8BUI6	Q8bui6 mus musculu
23	437	6.8	347	11	Q9D5M0	Q9d5m0 mus musculu
24	437	6.8	347	11	Q9D5D4	Q9d5d4 mus musculu
25	428.5	6.6	1416	4	Q9BZF9	Q9bzf9 homo sapien
26	428.5	6.6	1449	6	Q9BG87	Q9bg87 bos taurus
27	424.5	6.6	1401	6	Q8HYY4	Q8hyy4 bos taurus
28	415.5	6.4	500	11	Q810C2	Q810c2 rattus norv
29	410.5	6.4	315	13	Q7T2P5	Q7t2p5 brachydanio
30	402.5	6.2	300	11	Q9EQ29	Q9eq29 mus musculu
31	400	6.2	485	11	Q9D5K0	Q9d5k0 mus musculu
32	395	6.1	382	11	Q9D4J9	Q9d4j9 mus musculu
33	392.5	6.1	1388	6	Q9GL21	Q9gl21 canis famil
34	384.5	6.0	337	11	Q9CUC5	Q9cuc5 mus musculu
35	384.5	6.0	1201	11	Q8CIA8	Q8cia8 mus musculu
36	377	5.8	980	4	Q9P0K7	Q9p0k7 homo sapien
37	374	5.8	983	4	Q7Z5I4	Q7z5i4 homo sapien
38	373	5.8	980	4	Q7Z733	Q7z733 homo sapien
39	373	5.8	989	4	Q9P2L2	Q9p2l2 homo sapien
40	366	5.7	979	11	Q9EP71	Q9ep71 mus musculu
41	359	5.6	2055	5	Q8T5C7	Q8t5c7 plasmodium
42	359	5.6	2055	5	Q8IHP3	Q8ihp3 plasmodium
43	357.5	5.5	2083	5	Q9N435	Q9n435 caenorhabdi
44	357.5	5.5	10578	5	Q8ISF5	Q8isf5 caenorhabdi
45	357.5	5.5	18519	5	Q8ISF6	Q8isf6 caenorhabdi

ALIGNMENTS

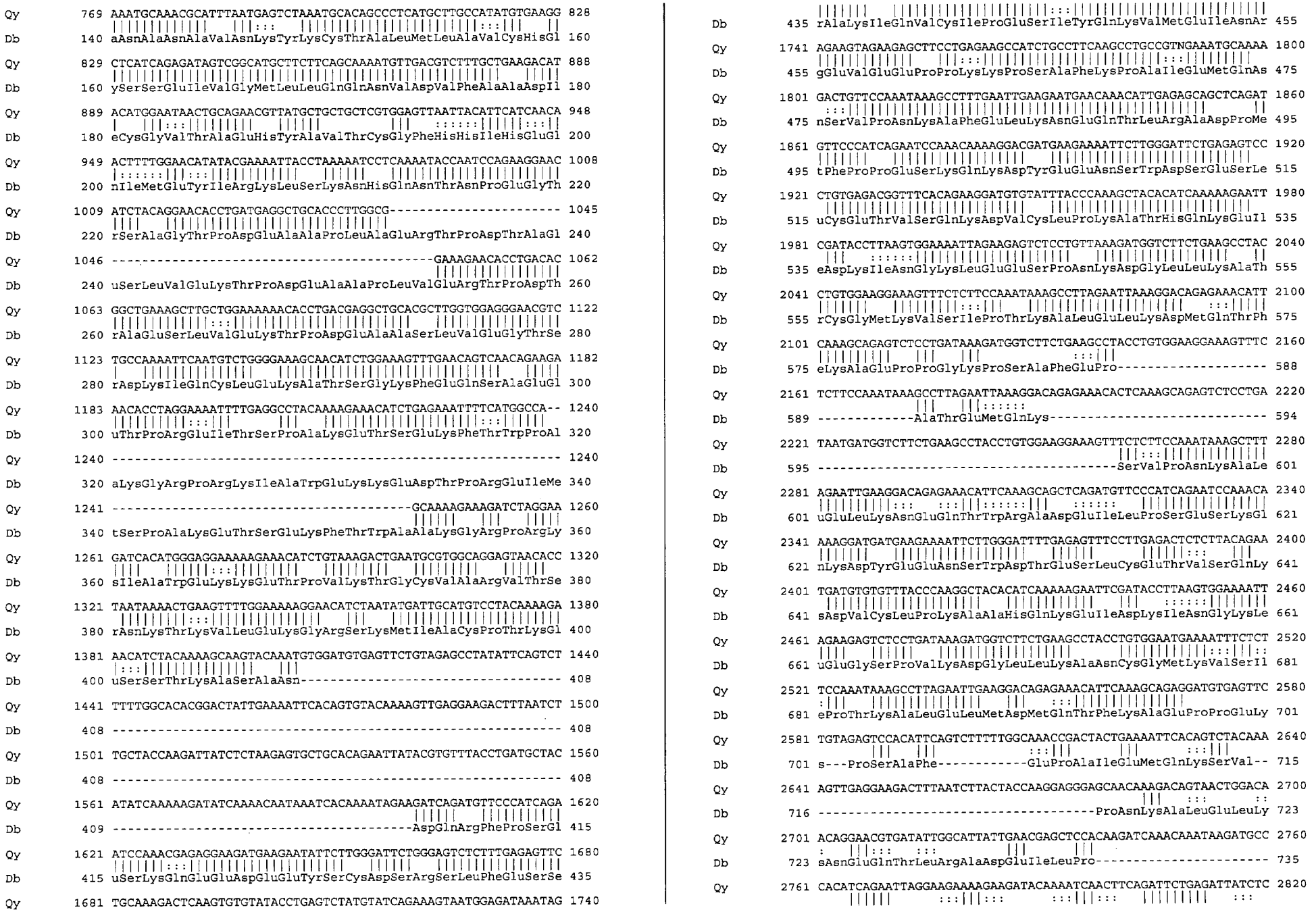
RESULT 1
Q9BXX2
ID Q9BXX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BXX2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269088; AAK27326.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 1.

```
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER    1011   1011
SQ   SEQUENCE   1011 AA;  114250 MW;  16627D7B218DE438 CRC64;

Alignment Scores:
Pred. No.:              3.56e-297      Length:          1011
Score:                  5173.00        Matches:         1011
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            80.23%         Indels:          0
DB:                     4              Gaps:            0

US-09-602-362E-26 (1-3673) x Q9BXX2 (1-1011)

Qy       641 ATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTC 700
Db         1 MetValAlaThrLeuLeuSerTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20

Qy       701 ACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACA 760
Db        21 ThrProLeuLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr 40

Qy       761 AAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATA 820
Db        41 LysAsnAlaAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle 60

Qy       821 TGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCT 880
Db        61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAla 80

Qy       881 GAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATT 940
Db        81 GluAspIleHisGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100

Qy       941 CATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCA 1000
Db       101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120

Qy      1001 GAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGAC 1060
Db       121 GluGlyThrSerThrGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAsp 140

Qy      1061 ACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACG 1120
Db       141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160

Qy      1121 TCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAA 1180
Db       161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180

Qy      1181 GAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA 1240
Db       181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerTrpPro 200

Qy      1241 GCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAA 1300
Db       201 AlaLysGluArgSerArgLysIleThrTrpGluGluLysGluThrSerValLysThrGlu 220

Qy      1301 TGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATG 1360
Db       221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240

Qy      1361 ATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCT 1420
Db       241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSer 260

Qy      1421 GTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAA 1480
Db       261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280

Qy      1481 GTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTAT 1540

Db       281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300

Qy      1541 ACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA 1600
Db       301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320

Qy      1601 GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCT 1660
Db       321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluGluTyrSerTrpAspSer 340

Qy      1661 GGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720
Db       341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360

Qy      1721 AAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAG 1780
Db       361 LysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380

Qy      1781 CCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAA 1840
Db       381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400

Qy      1841 ACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAAT 1900
Db       401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAspGluGluAsn 420

Qy      1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA 1960
Db       421 SerTrpAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440

Qy      1961 GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAA 2020
Db       441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluGluSerProValLys 460

Qy      2021 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2080
Db       461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480

Qy      2081 TTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCT 2140
Db       481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500

Qy      2141 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2200
Db       501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520

Qy      2201 CTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260
Db       521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540

Qy      2261 TCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320
Db       541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaAlaGlnMet 560

Qy      2321 TTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTC 2380
Db       561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580

Qy      2381 CTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTC 2440
Db       581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600

Qy      2441 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2500
Db       601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620

Qy      2501 TGTGGAATGAAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTC 2560
Db       621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640

Qy      2561 AAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACT 2620
Db       641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660
```

```
Qy       2621 GAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCA 2680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680

Qy       2681 ACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAA 2740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGln 700

Qy       2741 GATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACT 2800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720

Qy       2801 TCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCT 2860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SerAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAla 740

Qy       2861 ACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCT 2920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 ThrTyrGlnLysGluIleLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760

Qy       2921 TCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGG 2980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780

Qy       2981 AAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTT 3040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800

Qy       3041 CCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAA 3100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820

Qy       3101 ATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAA 3160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840

Qy       3161 ATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA 3220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IleLysSerGlnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860

Qy       3221 TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT 3280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 LeuProLeuAsnGlnGluGluGluLysArgArgAsnValAspIleLeuLysGluLysIle 880

Qy       3281 AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTC 3340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGluGlnThrLeu 900

Qy       3341 AGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACT 3400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920

Qy       3401 CATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATG 3460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940

Qy       3461 CTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTT 3520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrPhe 960

Qy       3521 GAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAA 3580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLys 980

Qy       3581 CAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCA 3640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 GlnLysThrValThrLysArgAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAla 1000

Qy       3641 GAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
             |||||||||||||||||||||||||||||||||
Db       1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011
```

```
RESULT 2
Q9BXX3
ID   Q9BXX3      PRELIMINARY;      PRT;  1341 AA.
AC   Q9BXX3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Breast cancer antigen NY-BR-1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21174979; PubMed=11280766;
RA   Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA   Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT   "Identification of a Tissue-specific Putative Transcription Factor in
RT   Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL   Cancer Res. 61:2055-2061(2001).
DR   EMBL; AF269087; AAK27325.1; -.
DR   GO; GO:0005634; C:nucleus; NAS.
DR   GO; GO:0005515; F:protein binding; NAS.
DR   GO; GO:0003700; F:transcription factor activity; NAS.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001969; Aspprotease_AS.
DR   Pfam; PF00023; ank; 6.
DR   SMART; SM00248; ANK; 6.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1341 AA;  152777 MW;  33E53DDE6FD3A58B CRC64;
```

```
Alignment Scores:
Pred. No.:              5.73e-196       Length:         1341
Score:                  3458.00         Matches:        740
Percent Similarity:     65.58%          Conservative:   83
Best Local Similarity:  58.96%          Mismatches:     159
Query Match:            53.63%          Indels:         273
DB:                     4               Gaps:           12

US-09-602-362E-26 (1-3673) x Q9BXX3 (1-1341)

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
              ||||||      ||||||   :::|||||||||   :::|||   :::||||||||||||
Db          1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
              ||||||||||||||||||||||||||      |||||||||||||||||||||||||||
Db         21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
              |||   |||:::|||||||||||||   |||||||||||||||||||||||||||||||
Db         41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
                 :::|||||| ||||||||||||||||||||:::|||||||||:::|||   |||||
Db         61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
              ||||||||||||    ||||||||||||||||||||||||    |||   :::|||||
Db         80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
              |   |||||||||:::|||||||||||||||||   |||||||||||||||||||||||
Db        100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
              |||||||:::|||   |||||||||:::|||   ||||||||||||||   ||||||||
Db        120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140
```



```
Db        736 ----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCy 754

Qy       2821 TGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAA 2880
                  :::   :::|||          |||||||||:::||||||:::||||||||||||
Db        754 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 774

Qy       2881 GACAACAAATGGCAAAATAGAAGAGTCT------------------------------ 2908
                     |||||||||:::|||||||||
Db        774 pLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCy 794

Qy       2908 ------------------------------------------------------------ 2908

Db        794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLy 814

Qy       2909 ----------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958
                       ||||||||||||||   ||||||||||   |||||||||   |||||
Db        814 sAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834

Qy       2959 TCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA-------------- 3004
                |||||||||   ||||||   ||||||:::||||||||||||||
Db        834 lProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854

Qy       3004 ------------------------------------------------------------ 3004

Db        854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874

Qy       3004 ------------------------------------------------------------ 3004

Db        874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894

Qy       3005 -------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTC 3045
                                ||||||||:::|||||||||||||||||   :::   ||
Db        894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914

Qy       3046 TTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGA 3105
                |||||||||   |||||||||:::||||||:::|||||||||   |||   ||||||||
Db        914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934

Qy       3106 ACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAA 3165
                |||||||||   ||||||||||||||:::|||:::|||||||||||||||||||||||
Db        934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954

Qy       3166 ATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCC 3225
                |||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db        954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974

Qy       3226 TTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285
                |||||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGl 994

Qy       3286 C---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGA 3330
                             |||||||||   ||||||:::||||||||||||   ||||||||
Db        994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014

Qy       3331 ACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGT 3390
                ||||   ||||||||||||||||||||||||||||||   |||||||||||||||||
Db       1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034

Qy       3391 TTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGA 3450
                |||||||||||||||:::||||||   |||   ||||||||||||||||||||||||
Db       1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

Qy       3451 AATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAA 3510
                |||||||||||||||||||:::|||||||||||||||||:::|||   ||||||||
Db       1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074

Qy       3511 TAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCT 3570
                |||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db       1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094

Qy       3571 AAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGT 3630
                |||||||||:::::::::::|||||||||||||||||||||   ||||||||||
Db       1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114
```

```
Qy       3631 TCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
                ||||   |||||||||||||||||||||||||||||||||
Db       1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128
```

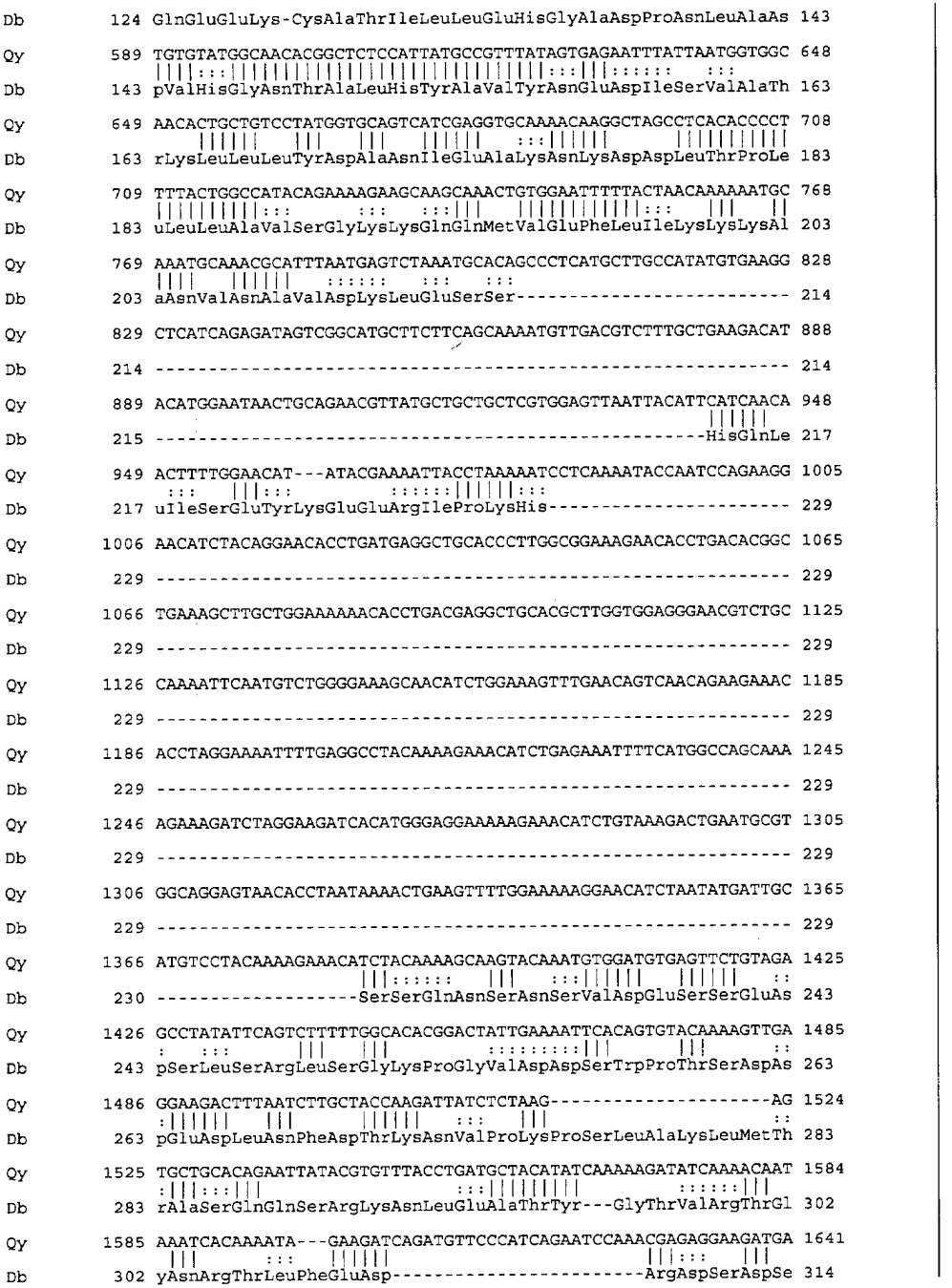
RESULT 3
Q9UPS8
ID Q9UPS8 PRELIMINARY; PRT; 1709 AA.
AC Q9UPS8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1074.
GN KIAA1074.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirosawa M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
DR EMBL; AB028997; BAA83026.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 4.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; ANK repeat; Repeat.
SQ SEQUENCE 1709 AA; 196351 MW; ADFEA9CDA067120E CRC64;

Alignment Scores:

Pred. No.:	4.36e-66	Length:	1709
Score:	1257.00	Matches:	383
Percent Similarity:	43.06%	Conservative:	166
Best Local Similarity:	30.04%	Mismatches:	323
Query Match:	19.49%	Indels:	403
DB:	4	Gaps:	33

US-09-602-362E-26 (1-3673) x Q9UPS8 (1-1709)

```
Qy        169 GGGGCTGCAGCCATGAAGAGGCTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAG 228
                 |||:::   |||   :::|||      :::|||   |||          |||   |||
Db         14 GlySerPheAlaArgArgArgArgSerSerAlaGlyGly-----------GlyGlyGlu 29

Qy        229 CCCCCGAACCCCTTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTC 288
                 |||               |||      |||::::::      |||      :::
Db         30 Pro------------GlyGluGlyAlaTyrSerGlnProGlyTyr---HisValArgAsp 44

Qy        289 GGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAG 348
                 |||||||||||||||||   |||||||||   |||   |||   |||:::::::::
Db         45 ArgAspLeuGlyLysIleHisLysAlaAlaSerAlaGlyAsnValAlaLysValGlnGln 64

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
                 :::   :::   :::|||          ||||||   ||||||          |||||||||
Db         65 IleLeuLeuLeuArgLysAsn---GlyLeuAsnAspArgAspLysMetAsnArgThrAla 83

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
                 ||||||   ||||||   |||||||||      |||||||||   ||||||||||||||
Db         84 LeuHisLeuAlaCysAlaAsnGlyHisProGluValValThrLeuLeuValAspArgLys 103

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
                 |||   |||||||||   |||   |||   ||||||   |||||||||||:::||||||
Db        104 CysGlnLeuAsnValCysAspAsnGluAsnArgThrAlaLeuMetLysAlaValGlnCys 123

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
                 :::   |||    ||||||   ||||||::::::   |||||||||   |||      ||
```


Db 124 GlnGluGluLys-CysAlaThrIleLeuLeuGluHisGlyAlaAspProAsnLeuAlaAs 143
Qy 589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
Db 143 pValHisGlyAsnThrAlaLeuHisTyrAlaValTyrAsnGluAspIleSerValAlaTh 163
Qy 649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
Db 163 rLysLeuLeuLeuTyrAspAlaAsnIleGluAlaLysAsnLysAspAspLeuThrProLe 183
Qy 709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
Db 183 uLeuLeuAlaValSerGlyLysLysGlnGlnMetValGluPheLeuIleLysLysLysAl 203
Qy 769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
Db 203 aAsnValAsnAlaValAspLysLeuGluSerSer---------- 214
Qy 829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
Db 214 ---------- 214
Qy 889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
Db 215 ----------HisGlnLe 217
Qy 949 ACTTTTGGAACAT---ATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGG 1005
Db 217 uIleSerGluTyrLysGluGluArgIleProLysHis---------- 229
Qy 1006 AACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGC 1065
Db 229 ---------- 229
Qy 1066 TGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGC 1125
Db 229 ---------- 229
Qy 1126 CAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAAC 1185
Db 229 ---------- 229
Qy 1186 ACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAA 1245
Db 229 ---------- 229
Qy 1246 AGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGT 1305
Db 229 ---------- 229
Qy 1306 GGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGC 1365
Db 229 ---------- 229
Qy 1366 ATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGA 1425
Db 230 ----------SerSerGlnAsnSerAsnSerValAspGluSerSerGluAs 243
Qy 1426 GCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGA 1485
Db 243 pSerLeuSerArgLeuSerGlyLysProGlyValAspAspSerTrpProThrSerAspAs 263
Qy 1486 GGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAG----------AG 1524
Db 263 pGluAspLeuAsnPheAspThrLysAsnValProLysProSerLeuAlaLysLeuMetTh 283
Qy 1525 TGCTGCACAGAATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAAT 1584
Db 283 rAlaSerGlnGlnSerArgLysAsnLeuGluAlaThrTyr---GlyThrValArgThrGl 302
Qy 1585 AAATCACAAAATA---GAAGATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGA 1641
Db 302 yAsnArgThrLeuPheGluAsp----------ArgAspSerAspSe 314
Qy 1642 AGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTAT 1701
Db 314 rGlnAspGluValValValGluSerLeuProThrThrSerIleLysValGln---CysPh 333
Qy 1702 ACCTGAGTCTATGTATCAG----------AAAGTAATGGA 1731
Db 333 eSerHisProThrTyrGlnSerProAspLeuLeuProLysProSerHisLysSerLeuAl 353
Qy 1732 GATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCT------GC 1785
Db 353 aAsnProGlyLeuMetLysGluGluProThrLysProGlyIleAlaLysLysGluAsnGl 373
Qy 1786 CGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATT 1845
Db 373 yIleAspIleIleGluSerAlaPro---------LeuGluGlnThrAsnAsnAspAsnLe 390
Qy 1846 G----------AGAGCAGCTCAGATGTTCCCATCAGAATC 1875
Db 390 uThrTyrValAspGluValHisLysAsnAsnArgSerAspMetMetSerAlaLeuGlyLe 410
Qy 1876 CAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTC 1935
Db 410 uGlyGlnGluGluAspIleGluSerProTrpAspSerGluSerIleSerGluAsnPhePr 430
Qy 1936 ACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGATACCTTAAGTGG 1995
Db 430 oGlnLysTyrVal----------AspProLeuAlaGl 439
Qy 1996 AAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGT 2055
Db 439 yAlaAlaAsp----------GlyLysGluLy 446
Qy 2056 TTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCC 2115
Db 446 sAsnIleGlyAsnGluGlnAlaGlu---------- 454
Qy 2116 TGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGC 2175
Db 455 -------AspValPheTyrIleProSerCys---MetSerGlySerArgAsnPheLysMe 471
Qy 2176 CTTAGAATTAAAGGACAGAGAAACACTC----------AAAGCAGAGTCTCCTGA 2220
Db 471 tAlaLysLeuGluAspThrArgAsnValGlyMetProValAlaHisMetGluSerProGl 491
Qy 2221 TAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTT 2280
Db 491 uArgTyrLeuHisLeuLysProThrIleGluMetLysAspSerValProAsnLysAlaGl 511
Qy 2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAG----------ATGTTCCC 2325
Db 511 yGlyMetLysAspValGlnThrSerLysAlaAlaGluHisAspLeuGluValAlaSerGl 531
Qy 2326 ATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGA 2385
Db 531 uGluGluGlnGluArgGluGlySerGluAsnAsnGlnProGlnValGluGluGluArgLy 551
Qy 2386 GACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATAC 2445
Db 551 sLysHisArgAsnAsnGluMetGluValSerAlaAsnIleHis----------AspGl 567
Qy 2446 CTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAG---------- 2494
Db 567 yAlaThrAspAspAlaGluAspAspAspAspAspGlyLeuIleGlnLysArgLysSe 587
Qy 2495 ----------CCTACCTGTGGAATGAAAATTTCTCTTCCAAATAA 2529
Db 587 rGlyGluThrAspHisGlnGlnPheProArgLysGluAsnLysGluTyrAlaSerGlyPr 607
Qy 2530 AGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTC 2589
Db 607 oAlaLeuGlnMetLysGluValLysSerThrGluLysGluLysArgThrSerLysGluSe 627

```
Qy       2590 CACATTCAGT---CTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGA 2646
Db        627 rValAsnSerProValPheGlyLysAlaSerLeuLeuThrGlyGlyLeuLeuGlnValAs 647

Qy       2647 GGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGA 2706
Db        647 pAspAspSerSerLeuSer--------------------------------GluIleAs 656

Qy       2707 ACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATC 2766
Db        656 pGluAspGluGlyArgProThrLysLysThrSerAsnGluLysAsnLysVal-------- 673

Qy       2767 AGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGAT---TCTGAGATTATCTCTGT 2823
Db        674 ------------LysAsnGlnIleGlnSerMetAspAspValAspAspLeuThrGlnSe 689

Qy       2824 GAGTGATACACAGAATTATGAGTGT---TTACCTGAGGCTACATATCAAAAAGAAATAAA 2880
Db        689 rSerGluThrAlaSerGluAspCysGluLeuProHisSerSerTyrLys----------- 705

Qy       2881 GACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCAC 2940
Db        705 ------------------------------------------------------------ 705

Qy       2941 TGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAG 3000
Db        706 -AsnPheMetLeuLeuIleGluGlnLeuGlyMetGluCysLys---------------- 719

Qy       3001 AGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAG 3060
Db        720 ----AspSerValSerLeuLeuLysIleGlnAspAlaAlaLeuSerCysGluArgLeuLe 738

Qy       3061 GGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAA 3120
Db        738 uGluLeuLysLysAsnHisCysGluLeuLeuThrValLysIleLysLysMetGluAspLy 758

Qy       3121 GTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAA 3180
Db        758 sValAsnValLeuGlnArgGluLeuSerGluThrLysGluIleLysSerGlnLeuGluHi 778

Qy       3181 CCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGA 3240
Db        778 sGlnLysValGluTrpGluArgGluLeuCysSerLeuArgPheSerLeuAsnGlnGluGl 798

Qy       3241 AGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT--------------AGACC 3285
Db        798 uGluLysArgArgAsnAlaAspThrLeuTyrGluLysIleArgGluGlnLeuArgArgLy 818

Qy       3286 CGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAAT 3345
Db        818 sGluGluGlnTyrArgLysGluValGluValLysGlnGlnLeuGluLeuSerLeuGlnTh 838

Qy       3346 ACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTT-------------- 3391
Db        838 rLeuGluMetGluLeuArgThrValLysSerAsnLeuAsnGlnValValGlnGluArgAs 858

Qy       3391 ------------------------------------------------------------ 3391
Db        858 nAspAlaGlnArgGlnLeuSerArgGluGlnAsnAlaArgMetLeuGlnAspGlyIleLe 878

Qy       3391 ------------------------------------------------------------ 3391
Db        878 uThrAsnHisLeuSerLysGlnLysGluIleGluMetAlaGlnLysLysMetAsnSerGl 898

Qy       3392 ----TCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAA 3447
Db        898 uAsnSerHisSerHisGluGluGluLysAspLeuSerHisLysAsnSerMetLeuGlnGl 918

Qy       3448 GGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGA 3507
Db        918 uGluIleAlaMetLeuArgLeuGluIleAspThrIleLysAsnGlnAsnGlnGluLysGl 938

Qy       3508 AAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGAC 3567
Db        938 uLysLysCysPheGluAspLeuLysIleValLysGluLysAsnGluAspLeuGlnLysTh 958

Qy       3568 CCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAA 3627
Db        958 rIleLysGlnAsnGluGluThrLeuThrGlnThrIleSerGlnTyrAsnGlyArgLeuSe 978

Qy       3628 AGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAG 3670
Db        978 rValLeuThrAlaGluAsnAlaMetLeuAsnSerLysLeuGlu 992
```

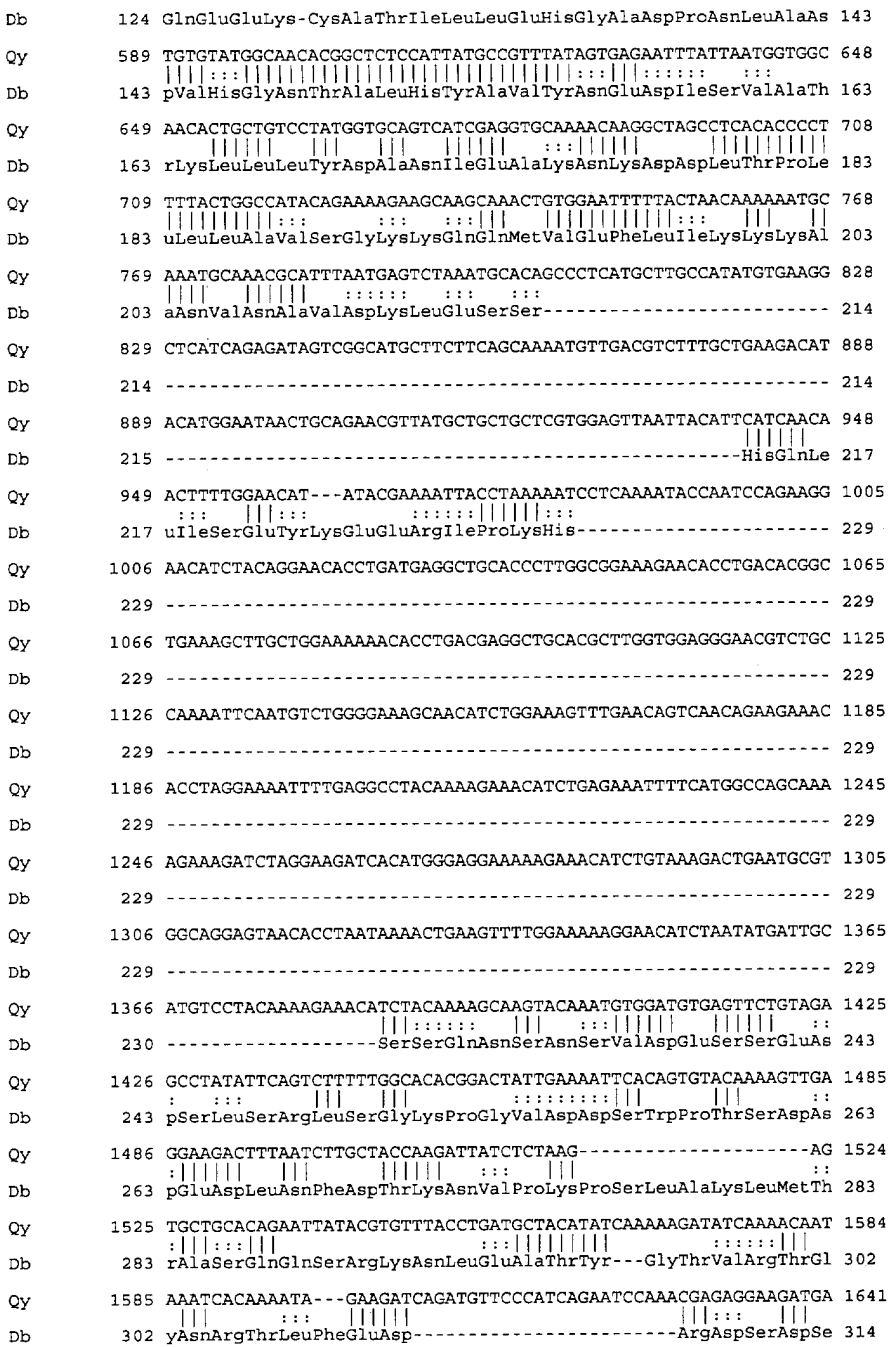
RESULT 4
Q9H1Q1
ID Q9H1Q1 PRELIMINARY; PRT; 1710 AA.
AC Q9H1Q1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BA145E8.1 (Hypothetical protein KIAA1074).
GN BA145E8.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 4.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; ANK repeat; Repeat.
SQ SEQUENCE 1710 AA; 196410 MW; 01CBF9BADB894872 CRC64;

Alignment Scores:

Pred. No.:	1.05e-65	Length:	1710
Score:	1250.50	Matches:	382
Percent Similarity:	43.10%	Conservative:	168
Best Local Similarity:	29.94%	Mismatches:	322
Query Match:	19.39%	Indels:	404
DB:	4	Gaps:	34

US-09-602-362E-26 (1-3673) x Q9H1Q1 (1-1710)

```
Qy        169 GGGGCTGCAGCCATGAAGAGGCTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAG 228
Db         14 GlySerPheAlaArgArgGlnArgSerSerAlaGlyGly------------GlyGlyGlu 29

Qy        229 CCCCCGAACCCCTTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTC 288
Db         30 Pro------------GlyGluGlyAlaTyrSerGlnProGlyTyr---HisValArgAsp 44

Qy        289 GGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAG 348
Db         45 ArgAspLeuGlyLysIleHisLysAlaAlaSerAlaGlyAsnValAlaLysValGlnGln 64

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
Db         65 IleLeuLeuLeuArgLysAsn---GlyLeuAsnAspArgAspLysMetAsnArgThrAla 83

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
Db         84 LeuHisLeuAlaCysAlaAsnGlyHisProGluValValThrLeuLeuValAspArgLys 103

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
Db        104 CysGlnLeuAsnValCysAspAsnGluAsnArgThrAlaLeuMetLysAlaValGlnCys 123

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
```


Db 124 GlnGluGluLys-CysAlaThrIleLeuLeuGluHisGlyAlaAspProAsnLeuAlaAs 143
Qy 589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
Db 143 pValHisGlyAsnThrAlaLeuHisTyrAlaValTyrAsnGluAspIleSerValAlaTh 163
Qy 649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
Db 163 rLysLeuLeuLeuTyrAspAlaAsnIleGluAlaLysAsnLysAspAspLeuThrProLe 183
Qy 709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
Db 183 uLeuLeuAlaValSerGlyLysLysGlnGlnMetValGluPheLeuIleLysLysLysAl 203
Qy 769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
Db 203 aAsnValAsnAlaValAspLysLeuGluSerSer---------- 214
Qy 829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
Db 214 ---------- 214
Qy 889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
Db 215 ----------HisGlnLe 217
Qy 949 ACTTTTGGAACAT---ATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGG 1005
Db 217 uIleSerGluTyrLysGluGluArgIleProLysHis---------- 229
Qy 1006 AACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGC 1065
Db 229 ---------- 229
Qy 1066 TGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGC 1125
Db 229 ---------- 229
Qy 1126 CAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAAC 1185
Db 229 ---------- 229
Qy 1186 ACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAA 1245
Db 229 ---------- 229
Qy 1246 AGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGT 1305
Db 229 ---------- 229
Qy 1306 GGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGC 1365
Db 229 ---------- 229
Qy 1366 ATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGA 1425
Db 230 ----------SerSerGlnAsnSerAsnSerValAspGluSerSerGluAs 243
Qy 1426 GCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGA 1485
Db 243 pSerLeuSerArgLeuSerGlyLysProGlyValAspAspSerTrpProThrSerAspAs 263
Qy 1486 GGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAG----------AG 1524
Db 263 pGluAspLeuAsnPheAspThrLysAsnValProLysProSerLeuAlaLysLeuMetTh 283
Qy 1525 TGCTGCACAGAATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAAT 1584
Db 283 rAlaSerGlnGlnSerArgLysAsnLeuGluAlaThrTyr---GlyThrValArgThrGl 302
Qy 1585 AAATCACAAAATA---GAAGATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGA 1641
Db 302 yAsnArgThrLeuPheGluAsp----------ArgAspSerAspSe 314
Qy 1642 AGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTAT 1701
Db 314 rGlnAspGluValValValGluSerLeuProThrThrSerIleLysValGln---CysPh 333
Qy 1702 ACCTGAGTCTATGTATCAG----------AAAGTAATGGA 1731
Db 333 eSerHisProThrTyrGlnSerProAspLeuLeuProLysProSerHisLysSerLeuAl 353
Qy 1732 GATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCT------GC 1785
Db 353 aAsnProGlyLeuMetLysGluGluProThrLysProGlyIleAlaLysLysGluAsnGl 373
Qy 1786 CGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATT 1845
Db 373 yIleAspIleIleGluSerAlaPro---------LeuGluGlnThrAsnAsnAspAsnLe 390
Qy 1846 G----------AGAGCAGCTCAGATGTTCCCATCAGAATC 1875
Db 390 uThrTyrValAspGluValHisLysAsnAsnArgSerAspMetMetSerAlaLeuGlyLe 410
Qy 1876 CAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTC 1935
Db 410 uGlyGlnGluGluAspIleGluSerProTrpAspSerGluSerIleSerGluAsnPhePr 430
Qy 1936 ACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGATACCTTAAGTGG 1995
Db 430 oGlnLysTyrVal----------AspProLeuAlaGl 439
Qy 1996 AAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGT 2055
Db 439 yAlaAlaAsp----------GlyLysGluLy 446
Qy 2056 TTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCC 2115
Db 446 sAsnIleGlyAsnGluGlnAlaGlu---------- 454
Qy 2116 TGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGC 2175
Db 455 -------AspValPheTyrIleProSerCys---MetSerGlySerArgAsnPheLysMe 471
Qy 2176 CTTAGAATTAAAGGACAGAGAAACACTC----------AAAGCAGAGTCTCCTGA 2220
Db 471 tAlaLysLeuGluAspThrArgAsnValGlyMetProValAlaHisMetGluSerProGl 491
Qy 2221 TAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTT 2280
Db 491 uArgTyrLeuHisLeuLysProThrIleGluMetLysAspSerValProAsnLysAlaGl 511
Qy 2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAG----------ATGTTCCC 2325
Db 511 yGlyMetLysAspValGlnThrSerLysAlaAlaGluHisAspLeuGluValAlaSerGl 531
Qy 2326 ATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGA 2385
Db 531 uGluGluGlnGluArgGluGlySerGluAsnAsnGlnProGlnValGluGluGluArgLy 551
Qy 2386 GACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATAC 2445
Db 551 sLysHisArgAsnAsnGluMetGluValSerAlaAsnIleHis----------AspGl 567
Qy 2446 CTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGG 2505
Db 567 yAlaThrAspAspAlaGluAspAspAspAspAspGlyLeuIleGlnLysArgLysSe 587
Qy 2506 AATGAAAATT---------TCTCTTCCAAATAAA---------- 2530
Db 587 rGlyGluThrAspHisGlnGlnPheProArgLysGluAsnLysGluTyrAlaSerSerGl 607
Qy 2531 ----GCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGA 2586
Db 607 yProAlaLeuGlnMetLysGluValLysSerThrGluLysGluLysArgThrSerLysGl 627

```
Qy       2587 GTCCACATTCAGT---CTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGT 2643
Db        627 uSerValAsnSerProValPheGlyLysAlaSerLeuLeuThrGlyGlyLeuLeuGlnVa 647

Qy       2644 TGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACA 2703
Db        647 lAspAspAspSerSerLeuSer----------------------------------GluIl 656

Qy       2704 GGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCAC 2763
Db        656 eAspGluAspGluGlyArgProThrLysLysThrSerAsnGluLysAsnLysVal----- 674

Qy       2764 ATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGAT---TCTGAGATTATCTC 2820
Db        675 ----------------LysAsnGlnIleGlnSerMetAspAspValAspAspLeuThrGl 689

Qy       2821 TGTGAGTGATACACAGAATTATGAGTGT---TTACCTGAGGCTACATATCAAAAAGAAAT 2877
Db        689 nSerSerGluThrAlaSerGluAspCysGluLeuProHisSerSerTyrLys-------- 706

Qy       2878 AAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGC 2937
Db        706 ------------------------------------------------------------ 706

Qy       2938 CACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATT 2997
Db        707 ----AsnPheMetLeuLeuIleGluGlnLeuGlyMetGluCysLys-------------- 720

Qy       2998 GAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGG 3057
Db        721 -------AspSerValSerLeuLeuLysIleGlnAspAlaAlaLeuSerCysGluArgLe 738

Qy       3058 AAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAA 3117
Db        738 uLeuGluLeuLysLysAsnHisCysGluLeuLeuThrValLysIleLysLysMetGluAs 758

Qy       3118 TAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGA 3177
Db        758 pLysValAsnValLeuGlnArgGluLeuSerGluThrLysGluIleLysSerGlnLeuGl 778

Qy       3178 GAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGA 3237
Db        778 uHisGlnLysValGluTrpGluArgGluLeuCysSerLeuArgPheSerLeuAsnGlnGl 798

Qy       3238 AGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT----------------AG 3282
Db        798 uGluGluLysArgArgAsnAlaAspThrLeuTyrGluLysIleArgGluGlnLeuArgAr 818

Qy       3283 ACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAG 3342
Db        818 gLysGluGluGlnTyrArgLysGluValGluValLysGlnGlnLeuGluLeuSerLeuGl 838

Qy       3343 AATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTT----------- 3391
Db        838 nThrLeuGluMetGluLeuArgThrValLysSerAsnLeuAsnGlnValValGlnGluAr 858

Qy       3391 ------------------------------------------------------------ 3391

Db        858 gAsnAspAlaGlnArgGlnLeuSerArgGluGlnAsnAlaArgMetLeuGlnAspGlyIl 878

Qy       3391 ------------------------------------------------------------ 3391

Db        878 eLeuThrAsnHisLeuSerLysGlnLysGluIleGluMetAlaGlnLysLysMetAsnSe 898

Qy       3392 -------TCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAA 3444
Db        898 rGluAsnSerHisSerHisGluGluGluLysAspLeuSerHisLysAsnSerMetLeuGl 918

Qy       3445 AAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAA 3504
Db        918 nGluGluIleAlaMetLeuArgLeuGluIleAspThrIleLysAsnGlnAsnGlnGluLy 938

Qy       3505 GGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAAT 3564
Db        938 sGluLysLysCysPheGluAspLeuLysIleValLysGluLysAsnGluAspLeuGlnLy 958

Qy       3565 GACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCT 3624
Db        958 sThrIleLysGlnAsnGluGluThrLeuThrGlnThrIleSerGlnTyrAsnGlyArgLe 978

Qy       3625 TAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAG 3670
Db        978 uSerValLeuThrAlaGluAsnAlaMetLeuAsnSerLysLeuGlu 993
```

RESULT 5
Q9NSI9
ID Q9NSI9 PRELIMINARY; PRT; 424 AA.
AC Q9NSI9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PRED4 protein (Fragment).
GN PRED4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.L.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL163203; CAB90394.1; -.
FT NON_TER 1 1
FT NON_TER 424 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

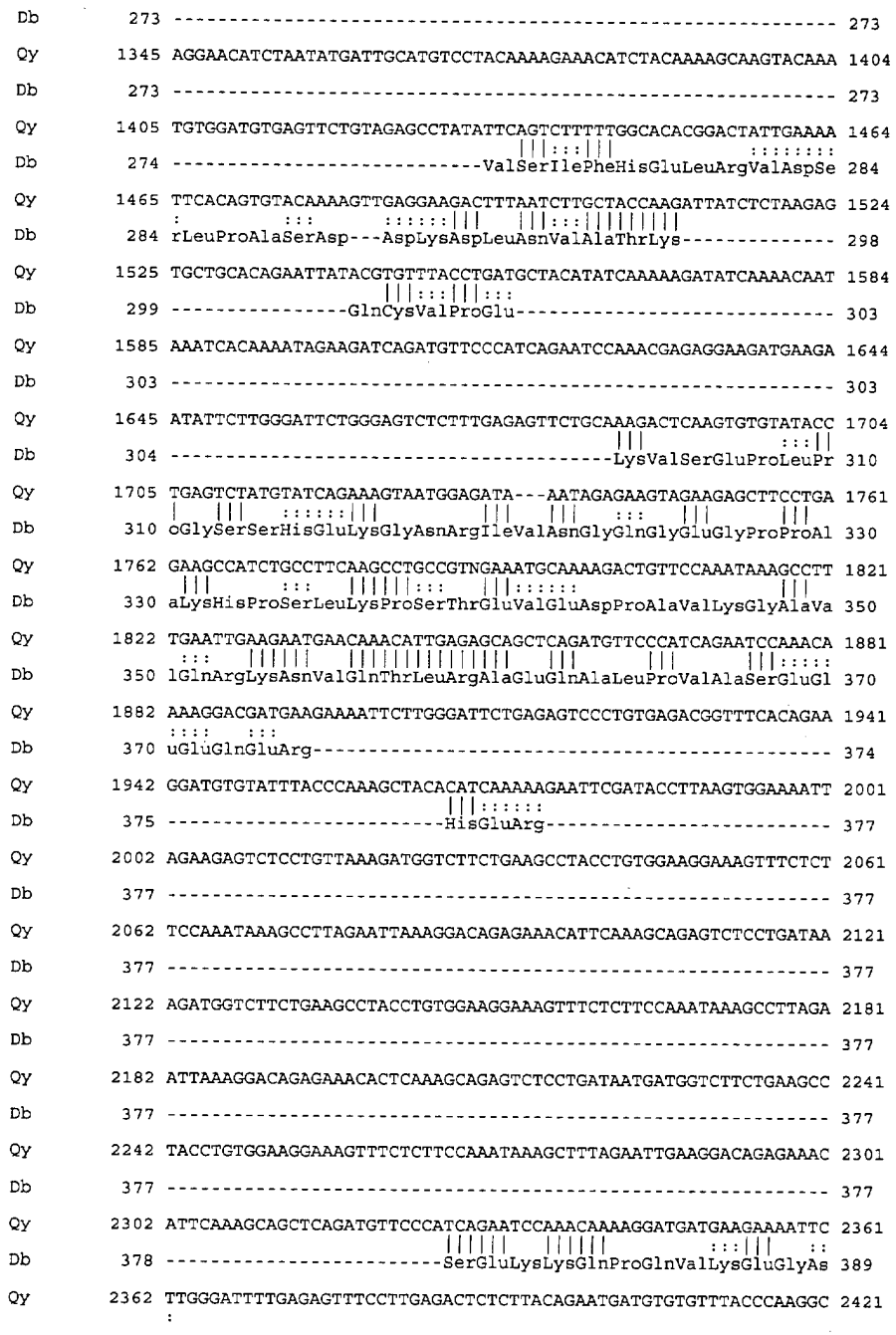
Alignment Scores:

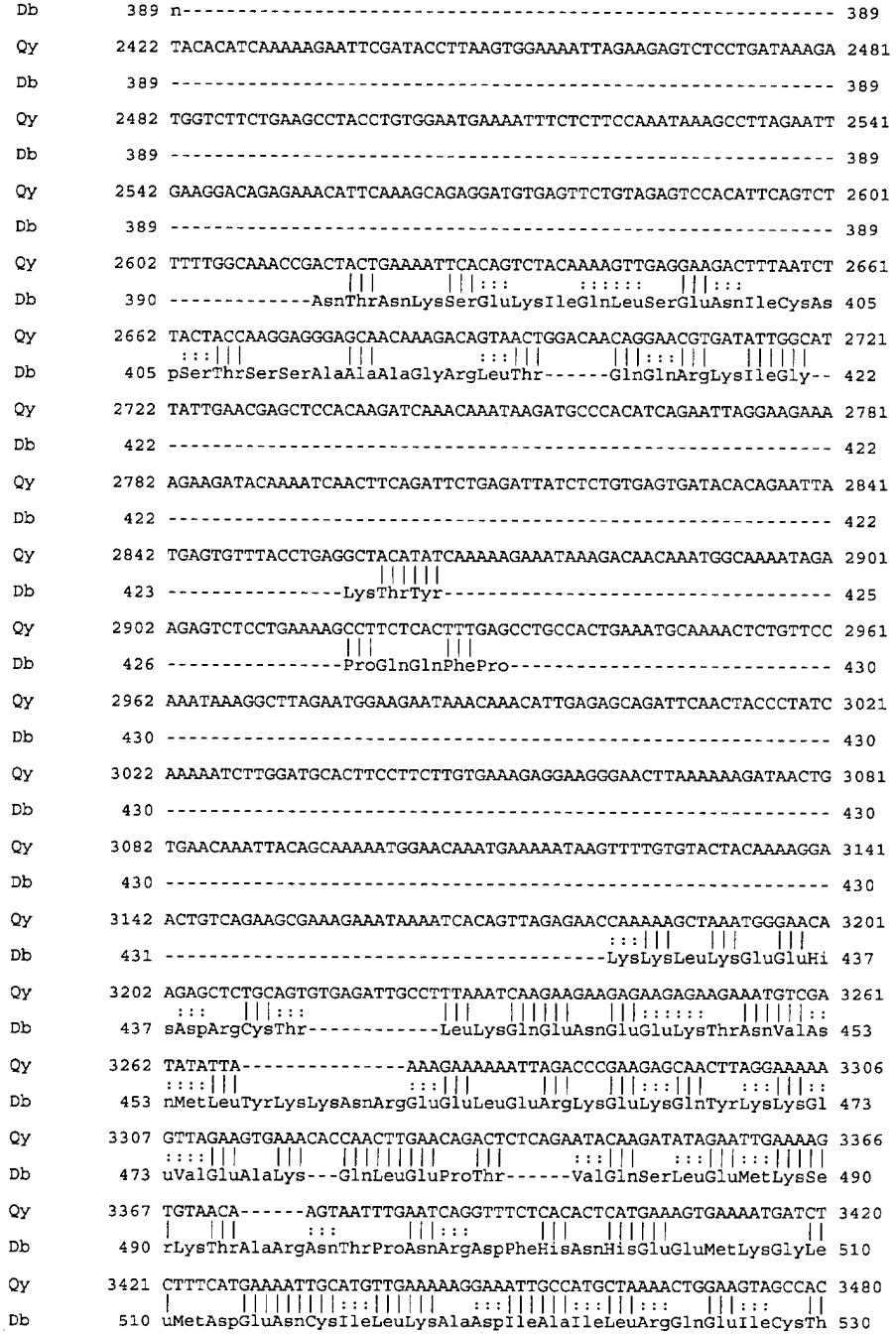
Pred. No.:	1.14e-51	Length:	424
Score:	1011.50	Matches:	204
Percent Similarity:	94.17%	Conservative:	6
Best Local Similarity:	91.48%	Mismatches:	4
Query Match:	15.69%	Indels:	9
DB:	4	Gaps:	1

US-09-602-362E-26 (1-3673) x Q9NSI9 (1-424)

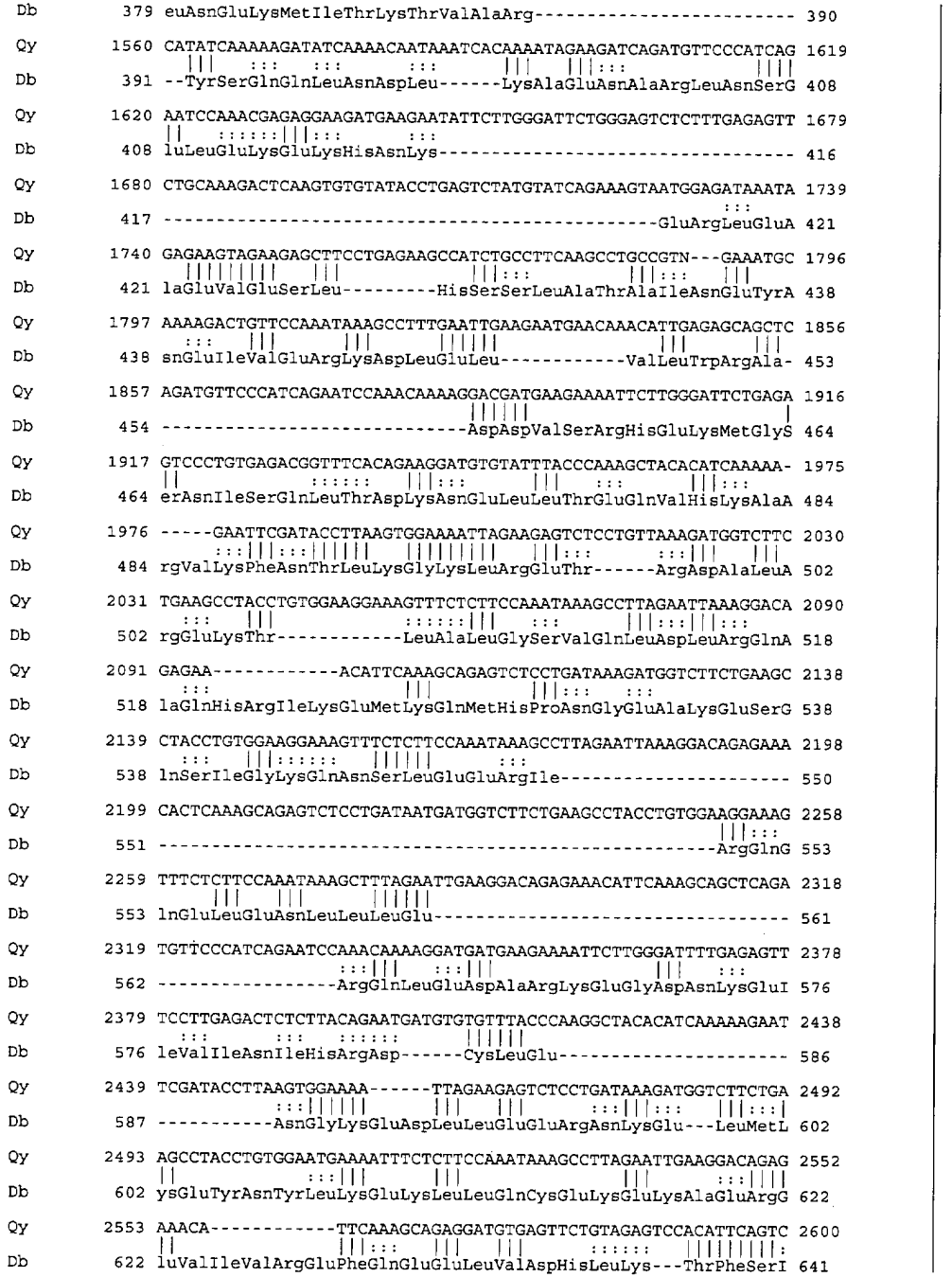
```
Qy       3005 GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAA 3064
Db          2 AspSerThrThrLeuSerLysIleLeuAspAlaValProSerCysGluArgGlyLysGlu 21

Qy       3065 CTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTT 3124
Db         22 LeuLysLysAspHisCysGluGlnIleThrAlaLysMetGluGlnThrLysAsnLysPhe 41

Qy       3125 TGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAA 3184
Db         42 CysValLeuGlnLysGluLeuSerGluAlaLysGluIleLysSerGlnLeuGluAsnGln 61

Qy       3185 AAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAG 3244
Db         62 LysAlaLysTrpGluGlnGluLeuCysSer-------------------------Lys 72

Qy       3245 AAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAA 3304
Db         73 LysArgArgAspValAspIleLeuLysGluLysIleArgProGluGluGlnLeuArgLys 92
```

```
Qy       3305 AAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAA 3364
Db         93 LysLeuGluValLysGlnGlnLeuGluGlnAlaLeuArgIleGlnAspIleGluLeuLys 112

Qy       3365 AGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTT 3424
Db         113 SerValThrSerAsnLeuAsnGlnValSerHisThrHisGluSerGluAsnAspLeuPhe 132

Qy       3425 CATGAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTG 3484
Db         133 HisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluValAlaThrLeu 152

Qy       3485 AAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAA 3544
Db         153 LysArgGlnHisGlnValLysGluAsnLysTyrPheGluAspIleLysIleLeuGlnGlu 172

Qy       3545 AAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCA 3604
Db         173 LysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGlnLysThrLeuThrLysArgAla 192

Qy       3605 TCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAA 3664
Db         193 SerGlnTyrArgGluGlnLeuLysValLeuThrAlaGluAsnThrMetLeuThrSerLys 212

Qy       3665 TTGAAGGAA 3673
Db         213 LeuLysGlu 215

RESULT 6
Q9H0H6
ID   Q9H0H6      PRELIMINARY;      PRT;   823 AA.
AC   Q9H0H6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   DKFZP434A171.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT   Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL   Genome Res. 11:422-435(2001).
DR   EMBL; AL136793; CAB66727.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 6.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   823 AA;  93968 MW;  9D2DDD5F52C379B2 CRC64;

Alignment Scores:
Pred. No.:              2.58e-33        Length:         823
Score:                  701.00          Matches:        245
Percent Similarity:     32.07%          Conservative:   128
Best Local Similarity:  21.07%          Mismatches:     207
Query Match:            10.87%          Indels:         583
DB:                     4               Gaps:           20

US-09-602-362E-26 (1-3673) x Q9H0H6 (1-823)

Qy        205 GGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCCTTCAGCGAACGGGTCTACACTGAG 264
Db          5 GlyPheGlySerArgArgGlyGlnThrAlaGlnGlySerIleAspHisValTyrThrGly 24

Qy        265 AAGGACTACGGGACCATCTACTTCGGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGG 324
Db         25 SerGlyTyr---ArgIleArgAspSerGluLeuGlnLysIleHisArgAlaAlaValLys 43

Qy        325 GGCCAAGTCCAGAAGCTGGAGAAGATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAA 384
Db         44 GlyAspAlaAlaGluValGluArg---CysLeuAlaArgArgSerGlyAspLeuAspAla 62

Qy        385 AGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAATGGCCATGCANAAGTA 444
Db         63 LeuAspLysGlnHisArgThrAlaLeuHisLeuAlaCysAlaSerGlyHisValGlnVal 82

Qy        445 GTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACA 504
Db         83 ValThrLeuLeuValAsnArgLysCysGlnIleAspValCysAspLysGluAsnArgThr 102

Qy        505 CCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGC 564
Db        103 ProLeuIleGlnAlaValHisCysGlnGluGluAla-CysAlaValIleLeuLeuGluHi 122

Qy        565 TGGTGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCTCCATTATGCCGTTTA 624
Db        122 sGlyAlaAsnProAsnPheLysAspIleTyrGlyAsnThrAlaLeuHisTyrAlaValTy 142

Qy        625 TAGTGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCA 684
Db        142 rSerGluSerThrSerLeuAlaGluLysLeuLeuSerHisGlyAlaHisIleGluAlaLe 162

Qy        685 AAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGT 744
Db        162 uAspLysAspAsnAsnThrProLeuLeuPheAlaIleIleCysLysLysGluLysMetVa 182

Qy        745 GGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGC 804
Db        182 lGluPheLeuLeuLysLysLysAlaSerSerHisAlaValAspArgLeuArgArgSerAl 202

Qy        805 CCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAA 864
Db        202 aLeuMetLeuAlaValTyrTyrAspSerProGlyIleValAsnIleLeuLeuLysGlnAs 222

Qy        865 TGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCG 924
Db        222 nIleAspValPheAlaGlnAspMetCysGlyArgAspAlaGluAspTyrAlaIleSerHi 242

Qy        925 TGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCC 984
Db        242 sHisLeuThrLysIleGlnGlnGlnIleLeuGluHisLysLysLysIleLeuLysLys-- 261

Qy        985 TCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGC 1044
Db        262 ----------------GluLysSerAspValGlySerSerAspGluSerAla-------- 273

Qy       1045 GGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACG 1104
Db        273 ------------------------------------------------------------ 273

Qy       1105 CTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTT 1164
Db        273 ------------------------------------------------------------ 273

Qy       1165 TGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGA 1224
Db        273 ------------------------------------------------------------ 273

Qy       1225 GAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAAC 1284
Db        273 ------------------------------------------------------------ 273

Qy       1285 ATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAA 1344
```





```
Qy       3481 ACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACA 3540
Db        530 rMetLysAsnAspAsnLeuGluLysGluAsnLysTyrLeuLysAspIleLysIleValLy 550

Qy       3541 AGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAG 3600
Db        550 sGluThrAsnAlaAlaLeuGluLysTyrIleLysLeuAsnGluGluMetIleThrGluTh 570

Qy       3601 GGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTC 3660
Db        570 rAlaPheArgTyrGlnGlnGluLeuAsnAspLeuLysAlaGluAsnThrArgLeuAsnAl 590

Qy       3661 TAAATTG 3667
Db        590 aGluLeu 592

RESULT 7
Q8IVF6
ID   Q8IVF6      PRELIMINARY;      PRT;   996 AA.
AC   Q8IVF6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein KIAA2015 (Fragment).
GN   KIAA2015.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Nagase T., Kikuno R., Ohara O.;
RT   "The nucleotide sequence of a long cDNA clone isolated from human.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB095935; BAC23111.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 6.
DR   SMART; SM00248; ANK; 6.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   996 AA;  116065 MW;  FD58060E3D8A1543 CRC64;

Alignment Scores:
Pred. No.:              1.06e-27      Length:          996
Score:                  606.00        Matches:         291
Percent Similarity:     40.86%        Conservative:    203
Best Local Similarity:  24.07%        Mismatches:      391
Query Match:            9.40%         Indels:          325
DB:                     4             Gaps:            53

US-09-602-362E-26 (1-3673) x Q8IVF6 (1-996)

Qy        190 CTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCCTTCAGCGAA 249
Db         19 LeuLeuSerSerMetAspGlnGluTyrAlaGlyProGly---------------------- 31

Qy        250 CGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTCGGGGATCTAGGGAAGATCCAT 309
Db         32 ------TyrAspIleArgAspTrp-----------------GluLeuArgLysIleHis 43

Qy        310 ACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAGATGACAGTAGGGAAGAAGCCC 369
Db         44 ArgAlaAlaIleLysGlyAspAlaAlaGluValGluArg---CysLeuThrArgArgPhe 62

Qy        370 GTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAAT 429
Db         63 ArgAspLeuAspAlaArgAspArgLysAspArgThrValLeuHisLeuAlaCysAlaHis 82

Qy        430 GGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGAT 489
Db         83 GlyArgValGlnValValThrLeuLeuLeuHisArgArgCysGlnIleAspIleCysAsp 102

Qy        490 GGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAA 549
Db        103 ArgLeuAsnArgThrProLeuMetLysAlaValHisSerGlnGluGluAla-CysAlaIl 122

Qy        550 TATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCT 609
Db        122 eValLeuLeuGluCysGlyAlaAsnProAsnIleLysAspIleTyrGlyAsnThrAlaLe 142

Qy        610 CCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGC 669
Db        142 uHisTyrAlaValTyrAsnLysGlyThrSerLeuAlaGluArgLeuLeuSerHisHisAl 162

Qy        670 AGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAG 729
Db        162 aAsnIleGluAlaLeuAsnLysGluGlyAsnThrProLeuLeuPheAlaIleAsnSerAr 182

Qy        730 AAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGA 789
Db        182 gArgGlnHisMetValGluPheLeuLeuLysAsnGlnAlaAsnIleHisAlaValAspAs 202

Qy        790 GTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCAT 849
Db        202 nPheLysArgThrAlaLeuIleLeuAlaValGlnHisAsnLeuSerSerIleValThrLe 222

Qy        850 GCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACG 909
Db        222 uLeuLeuGlnGlnAsnIleArgIleSerSerGlnAspMetPheGlyGlnThrAlaGluAs 242

Qy        910 TTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAA 969
Db        242 pTyrAlaLeuCysSerAspLeuArgSerIleArgGlnGlnIleLeuGluHisLysAsnLy 262

Qy        970 ATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGA 1029
Db        262 sMetLeuLysAsnHisLeuArgAsnAspAsnGlnGluThrAlaAlaMetLysPro-AlaA 282

Qy       1030 GGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACC 1089
Db        282 snLeu------------LysLysArgLysGluArgAlaLysAla---------------- 292

Qy       1090 TGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGC 1149
Db        293 ----------------------------------------------------------G 293

Qy       1150 AACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTAC 1209
Db        293 luHisAsnLeuLysValAlaSerGluGluLys---------------------------- 303

Qy       1210 AAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATCACATG 1269
Db        304 --GlnGluArgLeuGlnArgSerGluAsnLysGlnProGlnAsp---SerGlnSerTyrG 322

Qy       1270 GGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAATAAAA- 1328
Db        322 lyLysLysLys-----------------AspAlaMetTyrGlyAsnPheMetLeuLysL 336

Qy       1329 -----CTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAAC 1383
Db        336 ysAspIleAlaMetLeuLysGluGlu-----------LeuTyrAlaIleLysAsnAspS 352

Qy       1384 ATCTACAAAAGCAAGTACAAATGTG----GATGTGAGTTCTGTAGAGCCTATATTCAGTC 1439
Db        352 erLeuArgLysGluLysLysTyrIleGlnGluIleLysSerIleThrGluIleAsnAlaA 372

Qy       1440 TTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATC 1499
Db        372 snPhe-----------------------------------GluLysSerValArgL 379

Qy       1500 TTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTA 1559
```


Db 379 euAsnGluLysMetIleThrLysThrValAlaArg 390
Qy 1560 CATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAG 1619
Db 391 --TyrSerGlnGlnLeuAsnAspLeu------LysAlaGluAsnAlaArgLeuAsnSerG 408
Qy 1620 AATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTT 1679
Db 408 luLeuGluLysGluLysHisAsnLys 416
Qy 1680 CTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATA 1739
Db 417 GluArgLeuGluA 421
Qy 1740 GAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTN---GAAATGC 1796
Db 421 laGluValGluSerLeu---------HisSerSerLeuAlaThrAlaIleAsnGluTyrA 438
Qy 1797 AAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTC 1856
Db 438 snGluIleValGluArgLysAspLeuGluLeu-----------ValLeuTrpArgAla- 453
Qy 1857 AGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGA 1916
Db 454 -----AspAspValSerArgHisGluLysMetGlyS 464
Qy 1917 GTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAA- 1975
Db 464 erAsnIleSerGlnLeuThrAspLysAsnGluLeuLeuThrGluGlnValHisLysAlaA 484
Qy 1976 -----GAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTC 2030
Db 484 rgValLysPheAsnThrLeuLysGlyLysLeuArgGluThr------ArgAspAlaLeuA 502
Qy 2031 TGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACA 2090
Db 502 rgGluLysThr-----------LeuAlaLeuGlySerValGlnLeuAspLeuArgGlnA 518
Qy 2091 GAGAA-----------ACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGC 2138
Db 518 laGlnHisArgIleLysGluMetLysGlnMetHisProAsnGlyGluAlaLysGluSerG 538
Qy 2139 CTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAA 2198
Db 538 lnSerIleGlyLysGlnAsnSerLeuGluGluArgIle 550
Qy 2199 CACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAG 2258
Db 551 ArgGlnG 553
Qy 2259 TTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGA 2318
Db 553 lnGluLeuGluAsnLeuLeuGlu 561
Qy 2319 TGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTT 2378
Db 562 ----------ArgGlnLeuGluAspAlaArgLysGluGlyAspAsnLysGluI 576
Qy 2379 TCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAAT 2438
Db 576 leValIleAsnIleHisArgAsp------CysLeuGlu 586
Qy 2439 TCGATACCTTAAGTGGAAAA------TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGA 2492
Db 587 -----AsnGlyLysGluAspLeuLeuGluGluArgAsnLysGlu---LeuMetL 602
Qy 2493 AGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAG 2552
Db 602 ysGluTyrAsnTyrLeuLysGluLysLeuLeuGlnCysGluLysGluLysAlaGluArgG 622
Qy 2553 AAACA-----------TTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTC 2600
Db 622 luValIleValArgGluPheGlnGluGluLeuValAspHisLeuLys---ThrPheSerI 641
Qy 2601 TTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATC 2660
Db 641 leSerGluSerProLeuGluGlyThrSerHisCys----------HisIleAsnL 656
Qy 2661 TTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCA 2720
Db 656 euAsnGlu----------ThrTrpThrSerLysLysLysLeuPheG 668
Qy 2721 TTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAA 2780
Db 668 lnValGluIleGlnProGluGlu----------LysH 677
Qy 2781 AAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATT 2840
Db 677 isGluGluPheArgLysLeu------PheGluLeuIleSerLeu---------LeuAsnT 692
Qy 2841 ATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAG 2900
Db 692 yrThrAla----------AspGlnIleArgLysLysAsnArgGluLeuG 705
Qy 2901 AAGAGTCTCCTGAAAAGCCTTCTCAC---------TTTGAGCCTGCCACTGAAATGCAAA 2951
Db 705 luGlu------GluAlaThrGlyTyrLysLysCysLeuGluMetThrIleAsnMetLeuA 723
Qy 2952 ACTCTGTTCCAAATAAAGGCTTAGAATGGAAG---------AATAAACAAACATTGAGAG 3002
Db 723 snAlaPheAlaAsnGluAspPheAsnCysHisGlyAspLeuAsnThrAspGlnLeuLysM 743
Qy 3003 CAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGG 3062
Db 743 etAsp---IleLeuPheLysLysLeuLysGlnLysPheAsnAspLeuValAlaGluLysG 762
Qy 3063 AACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGT 3122
Db 762 luAlaValSerSerGluCysValAsnLeuAlaLysAspAsnGlu---------- 776
Qy 3123 TTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA---------TCACAGT 3173
Db 777 -----ValLeuHisGlnGluLeuLeuSerMetArgAsnValGlnGluLysCysGluLysL 795
Qy 3174 TAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA---------- 3220
Db 795 euGluLysAspLysLysMetLeuGluGluGluValLeuAsnLeuLysThrHisMetGluL 815
Qy 3221 --------TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGAT------ATAT 3266
Db 815 ysAspMetValGluLeuGlyLysLeuGlnGluTyrLysSerGluLeuAspGluArgAlaV 835
Qy 3267 TAAAAGAAAAAATTAGACCCGAAGAG---CAACTTAGGAAAAAGTTAGAAGTGAAACACC 3323
Db 835 alGlnGluIleGluLysLeuGluGluIleHisLeuGlnLysGlnAlaGluTyrGluLysG 855
Qy 3324 AACTTGAACAG----------ACTCTCAGAATACAAGATATAGAATTGA 3362
Db 855 lnLeuGluGlnLeuAsnLysAspAsnThrAlaSerLeuLysLysLysGluLeuThrLeuL 875
Qy 3363 AAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAA---------AGTGAAA 3413
Db 875 ysAspValGluCysLysPheSerLysMetLysThrAlaTyrGluGluValThrThrGluL 895
Qy 3414 ATGATCTCTTTCATGAA----------AATTGCATGTTGAAA- 3445
Db 895 euGluGluPheLysGluValPheAlaAlaAlaValLysAlaAsnAsnSerMetSerLysL 915
Qy 3446 ----------AAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAAC 3488
Db 915 ysLeuMetLysSerAspLysLysIleAlaValIleSerThrLysLeuPheThrGluLys- 934
Qy 3489 ATCAACACCAGGTGAAGGAAAATAAATACTTT----------G 3521
Db 935 ----------GlnArgMetLysTyrPheLeuSerThrLeuProThrArgProGluP 950

```
Qy       3522 AGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAAC 3581
Db        950 roGluLeuProCysValGluAsnLeuAsn--------------SerIleGluLeuAsnA 965

Qy       3582 AGAAAACAGTAACAAAAAGGGCA 3604
Db        965 rgLysTyrIleProLysThrAla 972

RESULT 8
Q9NW69
ID   Q9NW69      PRELIMINARY;      PRT;   546 AA.
AC   Q9NW69;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK001137; BAA91516.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 4.
DR   SMART; SM00248; ANK; 4.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER     546     546
SQ   SEQUENCE   546 AA;  60236 MW;  BEA458F57A917F83 CRC64;

Alignment Scores:
Pred. No.:              2.75e-27        Length:         546
Score:                  598.50          Matches:        186
Percent Similarity:     43.01%          Conservative:   91
Best Local Similarity:  28.88%          Mismatches:     187
Query Match:            9.28%           Indels:         180
DB:                     4               Gaps:           21

US-09-602-362E-26 (1-3673) x Q9NW69 (1-546)

Qy        169 GGGGCTGCAGCCATGAAGAGGCTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAG 228
Db         14 GlySerPheAlaArgArgArgArgSerSerAlaGlyGly-----------GlyGlyGlu 29

Qy        229 CCCCCGAACCCCTTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTC 288
Db         30 Pro------------GlyGluGlyAlaTyrSerGlnProGlyTyr---HisValArgAsp 44

Qy        289 GGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAG 348
Db         45 ArgAspLeuGlyLysIleHisLysAlaAlaSerAlaGlyAsnValAlaLysValGlnGln 64

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
Db         65 IleLeuLeuLeuArgLysAsn---GlyLeuAsnAspArgAspLysMetAsnArgThrAla 83

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
Db         84 LeuHisLeuAlaCysAlaAsnGlyHisProGluValValThrLeuLeuValAspArgLys 103

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
Db        104 CysGlnLeuAsnValCysAspAsnGluAsnArgThrAlaLeuMetLysAlaValGlnCys 123

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
Db        124 GlnGluGluLys-CysAlaThrIleLeuLeuGluHisGlyAlaAspProAsnLeuAlaAs 143

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
Db        143 pValHisGlyAsnThrAlaLeuHisTyrAlaValTyrAsnGluAspIleSerValAlaTh 163

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
Db        163 rLysLeuLeuLeuTyrAspAlaAsnIleGluAlaLysAsnLysAspAspLeuThrProLe 183

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
Db        183 uLeuLeuAlaValSerGlyLysLysGlnGlnMetValGluPheLeuIleLysLysLysAl 203

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
Db        203 aAsnValAsnAlaValAspLys---------------------------------------- 210

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
Db        211 ----------------------------------------------LeuGluSerSe 214

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
Db        214 rHisGlnLeuIleSerGlu--------------------------TyrLysGluGluAr 225

Qy        949 ACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAAC 1008
Db        225 gIleProLysHis-----------------SerSerGlnAsnSerAsnSerValAspGl 239

Qy       1009 ATCT----------------------ACAGGAACACCT------GATGAGGCTGCACC 1038
Db        239 uSerSerGluAspSerLeuSerArgLeuSerGlyLysProGlyValAspAspSerTrpPr 259

Qy       1039 CTTGGCGGAA--------------------------AGAACACCTGACACGGCTGAAAG 1071
Db        259 oThrSerAspAspGluAspLeuAsnPheAspThrLysAsnValProLysProSerLeuAl 279

Qy       1072 CTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAAT 1131
Db        279 aLysLeuMetThrAlaSerGlnGlnSerArgLysAsnLeuGluAlaThrTyrGlyThrVa 299

Qy       1132 TCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAG 1191
Db        299 lArg---ThrGlyAsnArgThr------LeuPheGluAspArgAspSerAspSerGlnAs 316

Qy       1192 GAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAG 1251
Db        316 pGluValValValGluSerLeuProThr------------------------------- 325

Qy       1252 ATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGG 1311
Db        326 ------------------------------ThrSerIleLysValGlnCysPheSerHi 335

Qy       1312 AGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGT-- 1369
Db        335 sProThrTyrGlnSerProAspLeuLeuProLysProSerHisLysSerLeuAlaAsnPr 355

Qy       1370 ------------------CCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGA 1410
Db        355 oGlyLeuMetLysGluGluProThrLysProGlyIleAlaLysLysGluAsnGlyIleAs 375

Qy       1411 TGTGAGTTCTGTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACA 1470
Db        375 pIle-----------------------------------------IleGluSerAlaPr 381

Qy       1471 GTGTACAAAAGTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGC 1530
```

```
Db        381 oLeuGluGlnThrAsnAsnAsp------------------------------------ 388

Qy       1531 ACAGAATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCA 1590
                                  :::   ||||||     ::::::        |||
Db        389 ----------------------AsnLeuThrTyrValAspGluValHisLysAsnAsnAr 401

Qy       1591 CAAAATAGAAGATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTC 1650
                    |||   |||                  :::|||||||||   |||
Db        401 g------SerAspMetMetSerAlaLeuGlyLeuGlyGlnGluGluAspIleGluSerPr 419

Qy       1651 TTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTG-------------- 1696
               |||||||||   |||:::   |||:::      :::      |||
Db        419 oTrpAspSerGluSerIleSerGluAsnPheProGlnLysTyrValAspProLeuAlaGl 439

Qy       1696 ------------------------------------------------------------ 1696

Db        439 yAlaAlaAspGlyLysGluLysAsnIleGlyAsnGluGlnAlaGluAspValPheTyrIl 459

Qy       1697 -------TGTATACCTGAGTCTATGTATCAGAAAGTA------ATGGAGATAAATAGAGA 1743
                    |||:::     |||        |||:::      :::|||      |||
Db        459 eProSerCysMetSerGlySerArgAsnPheLysMetAlaLysLeuGluAspThrArgAs 479

Qy       1744 AGTA-----------------GAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGC 1785
                |||                  |||   ||||||:::                |||||
Db        479 nValGlyMetProValAlaHisMetGluSerProGluArgTyrLeuHisLeuLysProTh 499

Qy       1786 CGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATT 1845
                :::||||||:::   :::|||||||||||||||      :::|||:::   ||||||
Db        499 rIleGluMetLysAspSerValProAsnLysAlaGlyGlyMetLysAspValGlnThrSe 519

Qy       1846 GAGAGCAGCT------------CAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGA 1893
                :::||||||            ::::::         |||   ::::::::::
Db        519 rLysAlaAlaGluHisAspLeuGluValAlaSerGluGluGlnGluArgGluGlySe 539

Qy       1894 AGAAAATTCT 1903
                ||||||:::
Db        539 rGluAsnAsn 542
```

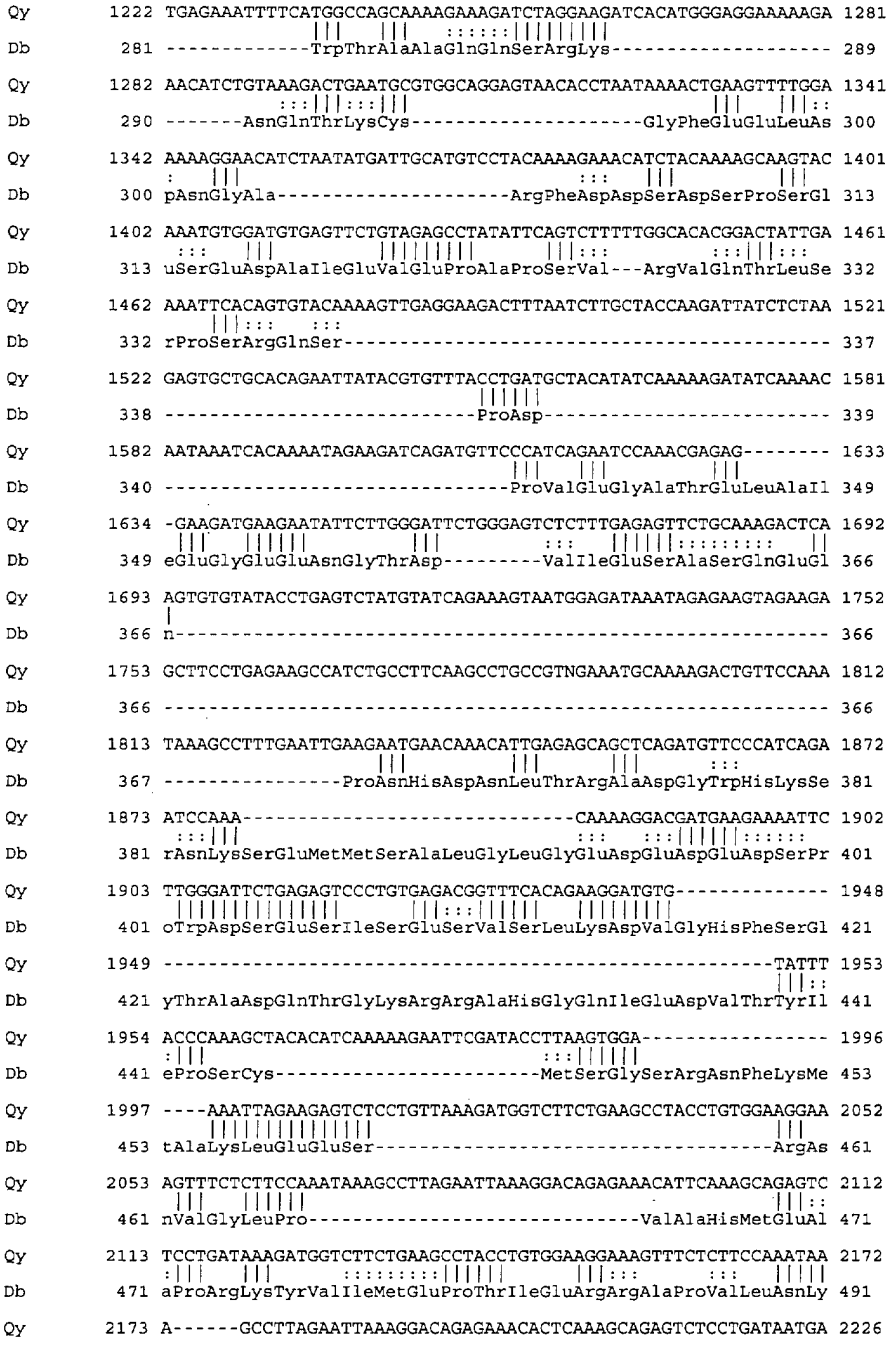
RESULT 9
Q811D2
ID   Q811D2   PRELIMINARY;   PRT;   733 AA.
AC   Q811D2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to RIKEN cDNA 5730521P14 gene (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II;
RA   Strausberg R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC047067; AAH47067.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 4.
DR   SMART; SM00248; ANK; 4.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
FT   NON_TER   733   733
SQ   SEQUENCE   733 AA; 80746 MW; 01FE3AAA1F69EA22 CRC64;

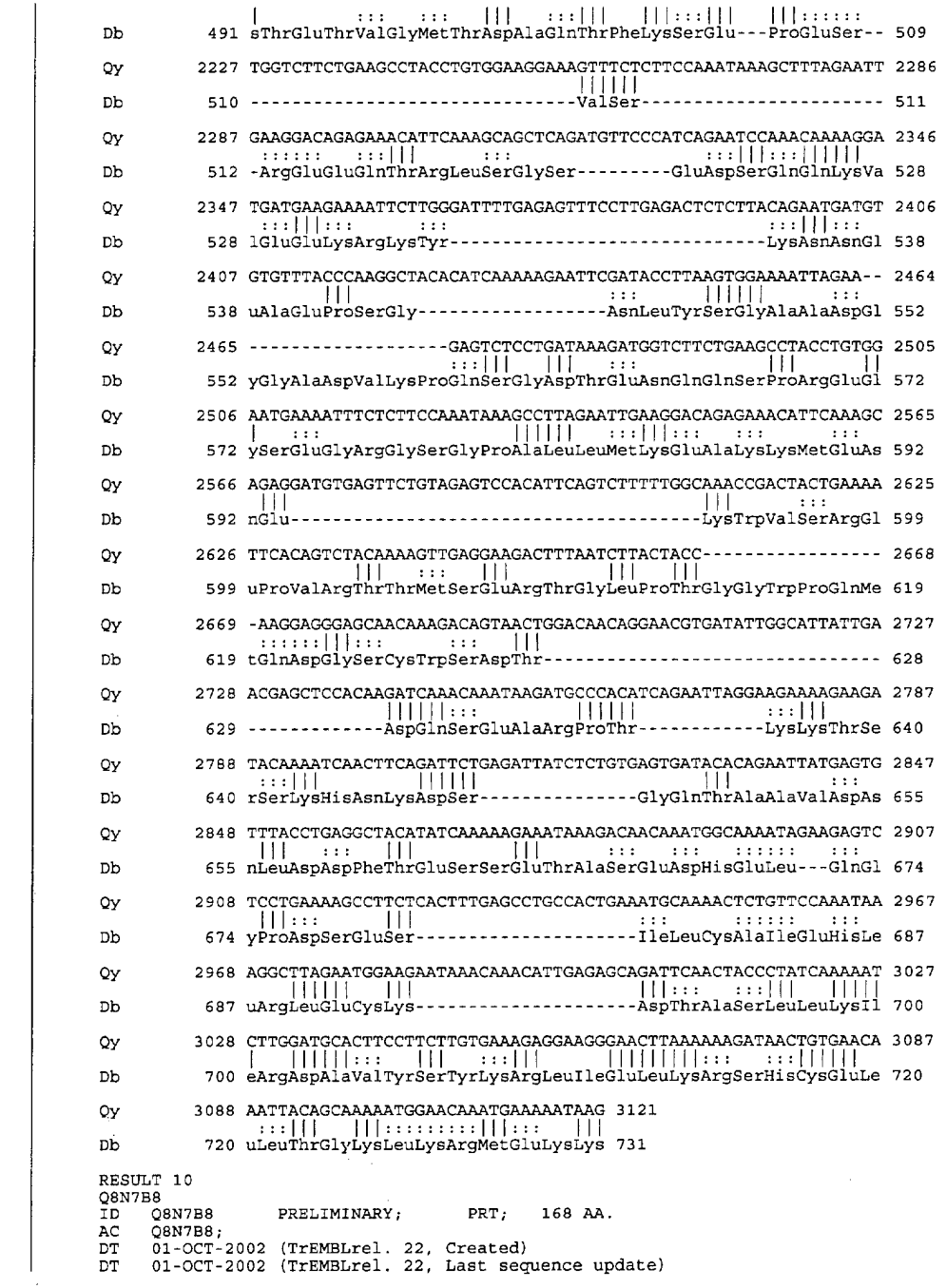
Alignment Scores:

Pred. No.:	2.2e-25	Length:	733
Score:	566.50	Matches:	251
Percent Similarity:	36.60%	Conservative:	134
Best Local Similarity:	23.86%	Mismatches:	275
Query Match:	8.79%	Indels:	392
DB:	11	Gaps:	41

US-09-602-362E-26 (1-3673) x Q811D2 (1-733)

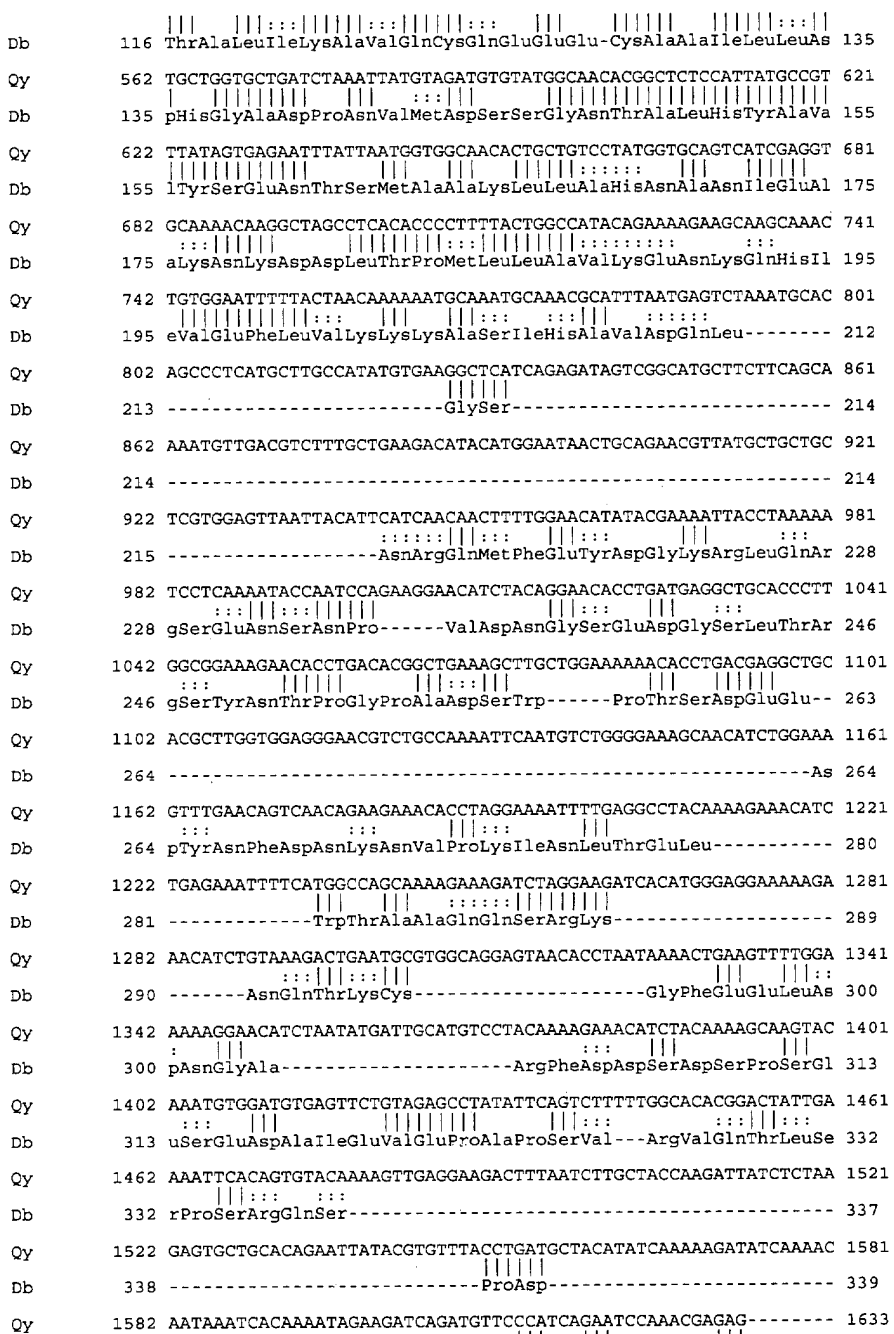
```
Qy        181 ATGAAGAGGCTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCC 240
              ||||||::::::              |||          ||||||   |||      |||
Db          1 MetLysLysIlePheGlyPheArgSerLys---------GlyProSerProLeuGlyPro 17

Qy        241 TTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACC--------------------- 279
                 :::                        :::|||
Db         18 SerAlaArgProArgSerAsnCysValGlyPheGlyArgGluSerAlaSerGlySerHis 37

Qy        280 --------------ATCTACTTCGGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGG 324
                           |||:::     |||:::||||||||||||   ||||||||
Db         38 ValProArgTyrHisIleHisAspLysAspMetGlyLysIleHisLysAlaAlaSerVal 57

Qy        325 GGCCAAGTCCAGAAGCTGGAGAAGATGACAGTA---GGGAAGAAGCCCGTCAACCTGAAC 381
              |||   |||   |||::::::   :::   :::   ||||||      |||      |||
Db         58 GlyAspValAlaLysValGlnHisIleLeuIleLeuGlyLysSerGlyVal------Asn 75

Qy        382 AAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAATGGCCATGCANAA 441
                  ||||||   |||   ||||||||||||||   ||||||      ||||||
Db         76 AspArgAspLysLysAspArgThrAlaLeuHisLeuAlaCysAlaTyrGlyHisProGlu 95

Qy        442 GTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGATGGCGAAGGGAGG 501
              ||||||||   ||||||:::|||||||||   ::::::      |||   |||
Db         96 ValValThrLeuLeuValGluArgLysCysGluIleAspAlaArgAspSerGluAsnSer 115

Qy        502 ACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAATATTCTCATAGA 561
              |||   |||:::||||||:::||||||:::   |||    ||||||   ||||||:::||
Db        116 ThrAlaLeuIleLysAlaValGlnCysGlnGluGluGlu-CysAlaAlaIleLeuLeuAs 135

Qy        562 TGCTGGTGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCTCCATTATGCCGT 621
              |   |||||||||   |||   :::|||      ||||||||||||||||||||||||||
Db        135 pHisGlyAlaAspProAsnAlaMetAspSerSerGlyAsnThrAlaLeuHisTyrAlaVa 155

Qy        622 TTATAGTGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGT 681
              ||||||||||||      |||   |||   ||||||::::::   |||   ||||||
Db        155 lTyrSerGluAsnThrSerMetAlaAlaLysLeuLeuAlaHisAsnAlaAsnIleGluAl 175

Qy        682 GCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAAC 741
               :::||||||      |||||||||:::||||||||:::::::::      :::
Db        175 aLysAsnLysAspAspLeuThrProMetLeuLeuAlaValLysGluAsnLysGlnHisIl 195

Qy        742 TGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCAC 801
              ||||||||||||:::   |||   |||:::   :::|||   ::::::
Db        195 eValGluPheLeuValLysLysLysAlaSerIleHisAlaValAspGlnLeu-------- 212

Qy        802 AGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCA 861
                                       ||||||
Db        213 -------------------------GlySer----------------------------- 214

Qy        862 AAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGC 921

Db        214 ------------------------------------------------------------ 214

Qy        922 TCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAA 981
                                       ::::::|||:::   |||:::      |||      :::
Db        215 -------------------AsnArgGlnIlePheGluTyrAspGlyLysArgLeuGlnAr 228

Qy        982 TCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTT 1041
                 :::|||:::||||||                  |||:::   |||   :::
Db        228 gSerGluAsnSerAsnPro------ValAspAsnGlySerGluAspGlySerLeuThrAr 246

Qy       1042 GGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGC 1101
               :::      ||||||     |||:::|||              |||   ||||||
Db        246 gSerTyrAsnThrProGlyProAlaAspSerTrp------ProThrSerAspGluGlu-- 263

Qy       1102 ACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAA 1161

Db        264 ----------------------------------------------------------As 264

Qy       1162 GTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATC 1221
               :::            :::      |||:::      |||
Db        264 pTyrAsnPheAspAsnLysAsnValProLysIleAsnLeuThrGluLeu---------- 280
```

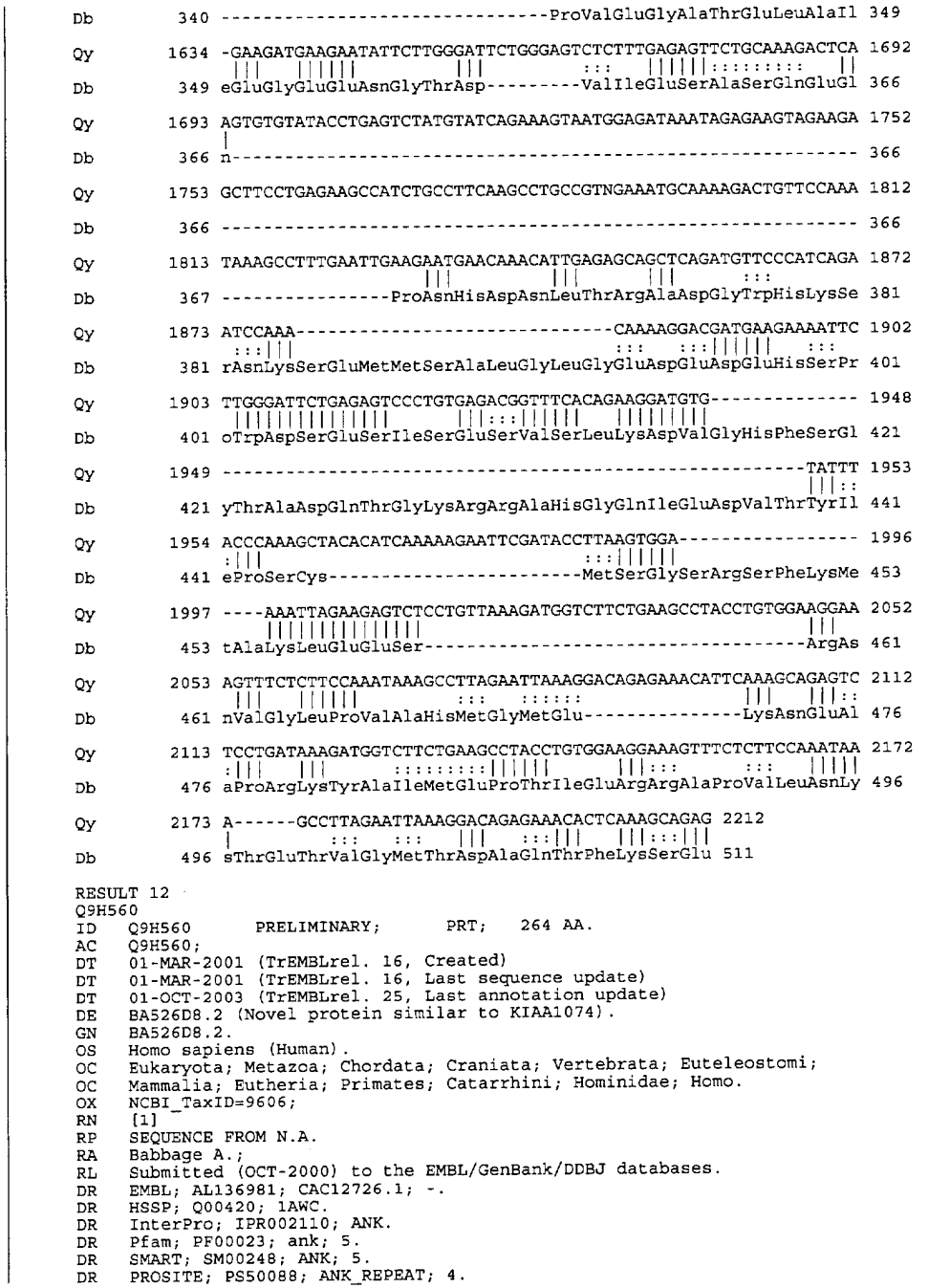




RESULT 10
Q8N7B8
ID Q8N7B8 PRELIMINARY; PRT; 168 AA.
AC Q8N7B8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

```
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ25822.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA   Arita M., Musashino K., Yuuki H., Hara H., Suzuki Y., Hata H.,
RA   Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA   Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA   Nagai K., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK098688; BAC05379.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   168 AA;  18305 MW;  BC6BA50DB58944CF CRC64;

Alignment Scores:
Pred. No.:              1.17e-24      Length:         168
Score:                  553.00        Matches:        113
Percent Similarity:     79.74%        Conservative:   9
Best Local Similarity:  73.86%        Mismatches:     27
Query Match:            8.58%         Indels:         4
DB:                     4             Gaps:           1

US-09-602-362E-26 (1-3673) x Q8N7B8 (1-168)

Qy       1064 GCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCT 1123
              ||||||||||||:::||||||   ||||||||||||||   ||||||||||||||:::
Db          2 AlaGluSerLeuValGluLysProProAspGluAlaAlaProLeuValGluGlyThrAla 21

Qy       1124 GCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAA 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         22 AspLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerAlaGluGlu 41

Qy       1184 ACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCA 1243
              ||||||:::||||||:::|||     ||||||||||||||||||||:::|||||||||
Db         42 ThrProLysLysIleMetArgThrAlaLysGluThrSerGluLysPheAlaTrpProAla 61

Qy       1244 AAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGC 1303
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db         62 LysGluArgProArgLysIleThrTrpGluGluLysGluThrSerValLysThrGluCys 81

Qy       1304 GTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATT 1363
              ||||||||||||   ||||||||||||||||||||||||||||||||||   |||:::
Db         82 ValAlaGlyValIleProAsnLysThrGluValLeuGluLysGlyThrSerGluMetLeu 101

Qy       1364 GCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTA 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 ThrCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSerVal 121

Qy       1424 GAGCCTATATTCAGTCTTTTT-----------GGCACACGGACTATTGAAAATTCACAG 1471
              |||   ||||||   :::           |||         :::      |||:::
Db        122 GluSerIlePheArgValSerProCysHisProGlyTrpAsnAlaMetAlaArgSerArg 141

Qy       1472 TGTACAAAAGTTGAGGAAGACTTTAATCTTGCTACCAAG 1510
              |||                 |||||||||
Db        142 LeuThrAlaThrSerThrSerGlnLysGluAlaThrLys 154

RESULT 11
Q9CS61
ID   Q9CS61      PRELIMINARY;      PRT;   533 AA.
AC   Q9CS61;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   5730521P14Rik protein (Fragment).
GN   5730521P14RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK017783; BAB30930.1; -.
DR   HSSP; P80144; 2MYO.
DR   MGD; MGI:1917887; 5730521P14Rik.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 4.
DR   SMART; SM00248; ANK; 4.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER     533     533
SQ   SEQUENCE   533 AA;  58545 MW;  FDB7E79116778430 CRC64;

Alignment Scores:
Pred. No.:              5.24e-22      Length:         533
Score:                  509.00        Matches:        191
Percent Similarity:     37.09%        Conservative:   82
Best Local Similarity:  25.95%        Mismatches:     180
Query Match:            7.89%         Indels:         283
DB:                     11            Gaps:           26

US-09-602-362E-26 (1-3673) x Q9CS61 (1-533)

Qy        181 ATGAAGAGGCTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCC 240
              ||||||::::::                ||| |||||| ||| |||
Db          1 MetLysLysIlePheGlyPheArgSerLys---------GlyProSerProLeuGlyPro 17

Qy        241 TTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACC---------- 279
                 :::                     :::|||
Db         18 SerAlaArgProArgSerAsnCysValGlyPheGlyArgGluSerAlaSerGlySerHis 37

Qy        280 ----------ATCTACTTCGGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGG 324
                       |||:::     |||:::|||||||||||   |||||||||
Db         38 ValProArgTyrHisIleHisAspLysAspMetGlyLysIleHisLysAlaAlaSerVal 57

Qy        325 GGCCAAGTCCAGAAGCTGGAGAAGATGACAGTA---GGGAAGAAGCCCGTCAACCTGAAC 381
              |||   |||   |||::::::   :::   :::   ||||||      |||    |||
Db         58 GlyAspValAlaLysValGlnHisIleLeuIleLeuGlyLysSerGlyVal------Asn 75

Qy        382 AAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAATGGCCATGCANAA 441
                 ||||||   |||   |||||||||||||   ||||||      ||||||
Db         76 AspArgAspLysLysAspArgThrAlaLeuHisLeuAlaCysAlaTyrGlyHisProGlu 95

Qy        442 GTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGATGGCGAAGGGAGG 501
              |||||||||   ||||||:::|||||||||   ::::::      |||   |||
Db         96 ValValThrLeuLeuValGluArgLysCysGluIleAspAlaArgAspSerGluSerSer 115

Qy        502 ACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAATATTCTCATAGA 561
```





RESULT 12
Q9H560
ID Q9H560 PRELIMINARY; PRT; 264 AA.
AC Q9H560;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BA526D8.2 (Novel protein similar to KIAA1074).
GN BA526D8.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL136981; CAC12726.1; -.
DR HSSP; Q00420; 1AWC.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 5.
DR SMART; SM00248; ANK; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.

```
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   264 AA;  30436 MW;  AB9237C94CAC0405 CRC64;

Alignment Scores:
Pred. No.:              5.19e-22        Length:         264
Score:                  508.50          Matches:        111
Percent Similarity:     62.50%          Conservative:   44
Best Local Similarity:  44.76%          Mismatches:     85
Query Match:            7.89%           Indels:         8
DB:                     4               Gaps:           2

US-09-602-362E-26 (1-3673) x Q9H560 (1-264)

Qy        241 TTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTCGGGGATCTAGGG 300
              ::::::   |||   |||      :::|||:::                  :::|||
Db         23 TyrAlaGlyArgGlyTyrHisIleArgAspTrp-----------------GluLeuArg 36

Qy        301 AAGATCCATACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAGATGACAGTAGGG 360
              |||||||||   ||||||   :::|||      :::::::|||                :::
Db         37 LysIleHisArgAlaAlaIleLysGlyAspAlaAlaGluValGluHis---CysLeuThr 55

Qy        361 AAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCC 420
              ::::::      :::|||:::   ||||||   |||   ||||||   ||||||
Db         56 ArgArgPheArgAspLeuAspAlaArgAspArgLysAspArgThrValLeuHisLeuThr 75

Qy        421 TGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAAT 480
              |||  :::|||          |||||||||   |||:::   |||:::|||   :::|||
Db         76 CysAlaHisGlyArgValGluValValThrLeuLeuLeuSerArgArgCysGlnIleAsn 95

Qy        481 GTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCT 540
              :::   |||          |||||||||||||||||||||:::   |||:::   ||||||
Db         96 IleTyrAspArgLeuAsnArgThrProLeuMetLysAlaValHisCysGlnGluGluAla 115

Qy        541 TTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGGCAA 600
              ||||||   ||||||::::::   ||||||:::   |||      |||:::|||   ||
Db        116 -CysAlaIleIleLeuLeuGluHisGlyAlaAsnProAsnIleLysAspIleTyrSerAs 135

Qy        601 CACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCTGTC 660
              |||||||||||||||||||||||||:::::::         :::         |||||||
Db        135 nThrAlaLeuHisTyrAlaValTyrAsnLysGlyThrSerLeuAlaGluLysLeuLeuSe 155

Qy        661 CTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCAT 720
              |:::   |||   ||||||      |||:::          ||||||||||||   |||||
Db        155 rHisHisAlaAsnIleGluAlaLeuAsnGluGluGlyAsnThrProLeuLeuPheAlaIl 175

Qy        721 ACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGC 780
              |      |||   :::|||   |||||||||||||||          ||||||   :::||
Db        175 eAsnSerArgArgGlnGlnIleValGluPheLeuLeuLysAsnGlnAlaAsnLeuHisAl 195

Qy        781 ATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGAT 840
              |   :::      :::   |||||||||||||||||:::          ||||||   ||
Db        195 aIleAspAsnPheArgArgThrAlaLeuMetLeuAlaValGlnHisAsnSerSerSerIl 215

Qy        841 AGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAAC 900
              ||||   :::|||||||||||||||:::::::::|||::::::|||:::   |||   ||
Db        215 eValSerLeuLeuLeuGlnGlnAsnIleAsnIlePheSerGlnAspLeuPheGlyGlnTh 235

Qy        901 TGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACA 960
              |||||||   ||||||                  |||   ||||||:::|||||||||
Db        235 rAlaGluAspTyrAlaValCysTyrAsnPheArgSerIleGlnGlnGlnIleLeuGluHi 255

Qy        961 TATACGAAAATTACCTAAAAAT 982
              |      |||:::   |||:::
Db        255 sLysAsnLysIleLeuLysSer 262

RESULT 13
Q7Z468
ID   Q7Z468      PRELIMINARY;      PRT;   357 AA.
AC   Q7Z468;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC056266; AAH56266.1; -.
KW   Hypothetical protein.
FT   NON_TER     357     357
SQ   SEQUENCE   357 AA;  41216 MW;  D6584C8C280ECA65 CRC64;

Alignment Scores:
Pred. No.:              1.95e-21        Length:         357
Score:                  499.00          Matches:        127
Percent Similarity:     54.65%          Conservative:   55
Best Local Similarity:  38.14%          Mismatches:     124
Query Match:            7.74%           Indels:         27
DB:                     4               Gaps:           6

US-09-602-362E-26 (1-3673) x Q7Z468 (1-357)

Qy        190 CTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCCTTCAGCGAA 249
              |||||:::::::      :::         ||||||
Db         15 LeuLeuSerSerMetAspGlnGluTyrAlaGlyProGly---------------------- 27

Qy        250 CGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTCGGGGATCTAGGGAAGATCCAT 309
                  |||      :::|||:::                  :::|||   |||||||||
Db         28 ------TyrAspIleArgAspTrp-----------------GluLeuArgLysIleHis 39

Qy        310 ACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAGATGACAGTAGGGAAGAAGCCC 369
                  ||||||   :::|||         ::::::|||:::      :::   ::::::
Db         40 ArgAlaAlaIleLysGlyAspAlaAlaGluValGluArg---CysLeuThrArgArgPhe 58

Qy        370 GTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAAT 429
              :::|||:::   ||||||   |||   ||||||   ||||||   ||||||   :::
Db         59 ArgAspLeuAspAlaArgAspArgLysAspArgThrValLeuHisLeuAlaCysAlaHis 78

Qy        430 GGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGAT 489
              |||          |||||||||   |||:::   |||:::|||   :::::::::   |||
Db         79 GlyArgValGlnValValThrLeuLeuLeuHisArgArgCysGlnIleAspIleCysAsp 98

Qy        490 GGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAA 549
                          |||||||||||||||||||||:::      :::   ||||||  ||||||
Db         99 ArgLeuAsnArgThrProLeuMetLysAlaValHisSerGlnGluGluAla-CysAlaIl 118
```

```
Qy       550 TATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCT 609
Db       118 eValLeuLeuGluCysGlyAlaAsnProAsnIleGluAspIleTyrGlyAsnThrAlaLe 138

Qy       610 CCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGC 669
Db       138 uHisTyrAlaValTyrAsnLysGlyThrSerLeuAlaGluArgLeuLeuSerHisHisAl 158

Qy       670 AGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAG 729
Db       158 aAsnIleGluAlaLeuAsnLysGluGlyAsnThrProLeuLeuPheAlaIleAsnSerAr 178

Qy       730 AAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGA 789
Db       178 gArgGlnHisMetValGluPheLeuLeuLysAsnGlnAlaAsnIleHisAlaValAspAs 198

Qy       790 GTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCAT 849
Db       198 nPheLysArgThrAlaLeuIleLeuAlaValGlnHisAsnLeuSerSerIleValThrLe 218

Qy       850 GCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACG 909
Db       218 uLeuLeuGlnGlnAsnIleArgIleSerSerGlnAspMetPheGlyGlnThrAlaGluAs 238

Qy       910 TTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAA 969
Db       238 pTyrAlaLeuCysSerAspLeuArgSerIleArgGlnGlnIleLeuGluHisLysAsnLy 258

Qy       970 ATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGA 1029
Db       258 sMetLeuLysAsnHisLeuArgAsnAspAsnGlnGluThrAlaAlaMetLysPro----- 276

Qy      1030 GGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACC 1089
Db       277 ----AlaAsnLeuLysLysArgLysGluArgAlaLysGluHisAsnLeuLysValAlaSe 295

Qy      1090 TGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAA------------ATTCAATG 1137
Db       295 rGluGluLysGlnGluArgLeuGlnArgSerGluAsnLysGlnProGlnAspSerGlnSe 315

Qy      1138 TCTGGGGAAA---------GCAACATCTGGAAAGTTT 1165
Db       315 rTyrGlyLysLysLysAspAlaMetTyrGlyAsnPhe 327
```

RESULT 14
Q9HCD1
ID Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1641 (Fragment).
GN KIAA1641.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046861; BAB13467.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Alignment Scores:
Pred. No.: 8.2e-21 Length: 718
Score: 489.00 Matches: 160

Percent Similarity: 42.60% Conservative: 79
Best Local Similarity: 28.52% Mismatches: 186
Query Match: 7.58% Indels: 136
DB: 4 Gaps: 12

US-09-602-362E-26 (1-3673) x Q9HCD1 (1-718)

```
Qy      2009 TCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAAT 2068
Db        11 SerSerGlnLysGlnProAlaLeuLysAlaThrSerAspLysLysAspSerValSerAsn 30

Qy      2069 AAAGCCTTAGAATTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGT 2128
Db        31 IleProThrGluIleLysAspGlyGlnGlnSerGlyThrValSerSerGlnLysGlnPro 50

Qy      2129 CTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAG 2188
Db        51 AlaTrpLysAlaThrSerValLysLysAspSerValSerAsnIleAlaThrGluIleLys 70

Qy      2189 GACAGAGAAACACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGT 2248
Db        71 AspGlyGlnIleArgGlyThrValSerProGlnLysGlnSerAlaGlnLysValIlePhe 90

Qy      2249 GGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAA 2308
Db        91 LysLysLysValSerLeuLeuAsnIleAlaThrArgIle---------------------- 103

Qy      2309 GCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGAT 2368
Db       104 ----------------------------------------------ThrGlyGlyTrpLys 108

Qy      2369 TTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACAT 2428
Db       108 ------------------------------------------------------------ 108

Qy      2429 CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTT 2488
Db       109 --------------------SerGly------ThrGluTyrProGluAsnLeuProThr 119

Qy      2489 CTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGAC 2548
Db       120 LeuLysAlaThrIleGluAsnLysAsnSerValLeuAsnThrAlaThrLysMetLysAsp 139

Qy      2549 AGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGC 2608
Db       139 ------------------------------------------------------------ 139

Qy      2609 AAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACC 2668
Db       140 -----------------ValGlnThrSerThrProGluGlnAspLeuGluMetAlaSer 153

Qy      2669 AAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAA 2728
Db       154 Glu----------------------GlyGluGlnLysArg----------------- 159

Qy      2729 CGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGAT 2788
Db       160 ------LeuGluGluTyrGluAsnAsnGlnProGlnVal-----------LysAsnGln 173

Qy      2789 ACAAAATCAACTTCAGAT---TCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAG 2845
Db       174 IleHisSerArgAspAspLeuAspAspIleIleGlnSerSerGlnThrValSer------ 191

Qy      2846 TGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAG 2905
Db       191 ------------------------------------------------------------ 191

Qy      2906 TCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAAT 2965
Db       192 ----------------------GluAspGlyAspSerLeuCysCysAsnCysLysAsn 203

Qy      2966 AAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAA 3025
```

```
Db        204 ValIleLeuLeuIleAspGlnHisGluMetLysCysLysAspCysValHisLeuLeuLys 223

Qy       3026 ATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAA 3085
                |||   :::                 :::|||      :::||||||   ::::::||||||
Db        224 IleLysAsnThrPheCysLeuTrpLysArgLeuIleLysLeuLysAspAsnHisCysGlu 243

Qy       3086 CAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTG 3145
                |||:::      |||:::   ::::::||||||||      ||||||||||   :::
Db        244 GlnLeuArgValLysIleArgLysLeuLysAsnLysAlaSerValLeuGlnLysArgIle 263

Qy       3146 TCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAG 3205
                ||||||   :::|||||||||||||||||||:::::::::      :::   |||:::|||
Db        264 SerGluLysGluGluIleLysSerGlnLeuLysHisGluIleLeuGluLeuGluLysGlu 283

Qy       3206 CTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATA 3265
                |||||||||:::|||      :::   ||||||::::::|||||||||||||||:::
Db        284 LeuCysSerLeuArgPheAlaIleGlnGlnGluLysLysLysArgArgAsnValGluGlu 303

Qy       3266 TTAAAAGAAAAAATTAGA--------------CCCGAAGAGCAACTTAGGAAAAAGTTA 3310
                |||   :::|||:::|||                ||||||||   |||   :::
Db        304 LeuHisGlnLysValArgGluLysLeuArgIleThrGluGluGlnTyrArgIleGluAla 323

Qy       3311 GAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTA 3370
                :::|||          ::::::      |||:::      ::::::||||||||:::
Db        324 AspValThrLysProIleLysProAlaLeuLysSerAlaGluValGluLeuLysThrGly 343

Qy       3371 ACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAA 3430
                :::|||   ||||||||||||   |||   |||   |||   ||||||   ||||||
Db        344 GlyAsnAsnSerAsnGlnValSerGluThrAspGluLysGlu---AspLeuLeuHisGlu 362

Qy       3431 AATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACAT 3490
                |||   :::::::::   |||||||||   |||:::||||||      ||:::|||:::
Db        363 AsnArgLeuMetGlnAspGluIleAlaArgLeuArgLeuGluLysAspThrIleLysAsn 382

Qy       3491 CAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAAT 3550
                |||:::          |||   ||||||   :::|||   :::|||::::::   |||:::
Db        383 GlnAsn------LeuGluLysLysTyrLeuLysAspPheGluIleValLysArgLysHis 400

Qy       3551 GCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAG 3610
                :::||||||      ||||||         :::|||:::   |||      :::
Db        401 GluAspLeuGlnLysAlaLeuLysArgAsnGlyGluThrLeuAlaLysThrIleAlaCys 420

Qy       3611 TATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAG 3670
                |||      ||||||      ||||||   |||||||||   |||   |||||||||:::
Db        421 TyrSerGlyGlnLeuAlaAlaLeuThrAspGluAsnThrThrLeuArgSerLysLeuGlu 440

Qy       3671 GAA 3673
                :::
Db        441 Lys 441
```

```
RESULT 15
Q9D504
ID   Q9D504      PRELIMINARY;      PRT;   279 AA.
AC   Q9D504;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   4930532L20Rik protein.
GN   4930532L20RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK015948; BAB30047.1; -.
DR   HSSP; P25963; 1IKN.
DR   MGD; MGI:1922446; 4930532L20Rik.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   279 AA;  31247 MW;  484FB28A84E36CA1 CRC64;
```

```
Alignment Scores:
Pred. No.:              9.69e-21        Length:         279
Score:                  487.00          Matches:        108
Percent Similarity:     60.94%          Conservative:   48
Best Local Similarity:  42.19%          Mismatches:     87
Query Match:            7.55%           Indels:         13
DB:                     11              Gaps:           3

US-09-602-362E-26 (1-3673) x Q9D504 (1-279)

Qy        229 CCCCCGAACCCCTTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTC 288
                |||||||||                  |||      :::|||
Db         32 ProProAsnLeuSerIleGlyGlyGlyTyrArgLeuArgAsp----------------- 45

Qy        289 GGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAG 348
                  |||   |||:::|||   ||||||:::   |||      |||||||||   |||
Db         46 LysHisLeuLysLysLeuHisLysAlaAlaThrIleGlyAsnGluGlnLysLeu---Lys 64

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
                  :::   :::|||      |||:::|||   ||||||   :::   ||||||
Db         65 AspTyrLeuGluArgLysLysTyrAsnValAsnGlyArgAspLysArgSerArgThrPro 84

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
                ||||||   ||||||   ||||||:::      :::|||:::   |||:::::   :::
Db         85 LeuHisLeuAlaCysAlaAsnGlyTyrThrAsnIleValSerLeuLeuIleGluAsnGln 104

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
                |||   :::||||||   |||   |||   ||||||||||||:::||||||::::::|||
Db        105 CysLysIleAsnValGlnAspSerGluAsnArgThrProLeuIleLysAlaValGluCys 124

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
                ::::::|||::: ||||||   :::|||:::      |||||||||   |||   |||||
Db        125 GlnGlnGluSer-CysAlaThrValLeuLeuLeuHisGlyAlaAspProAsnLeuValAs 144

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
                |||||||   ||||||||||||||||||||||||      :::|||:::   :::
Db        144 pValTyrSerAsnThrAlaLeuHisTyrAlaValCysGlyGlnAsnIleSerLeuAlaAs 164

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
                  ||||||   |||   |||   :::|||   :::||||||         ||||||||
Db        164 nLysLeuLeuGlnTyrLysAlaAsnLeuGluAlaLysAsnLysAspGlyHisThrProLe 184

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
                ||||||||||:::   :::   ::::::      |||:::|||||||||   |||   ||
Db        184 uLeuLeuAlaValAlaGluAsnAsnGluAsnMetValLysPheLeuLeuLysLysGlyAl 204

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
```

```
                |:::   ||||||   :::::::::      ||||||:::|||:::|||:::
Db        204 aAspValAsnAlaSerAspLysAsnHisArgThrAlaIleMetIleAlaLeuIleValGl 224

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
                :::      |||   :::|||||||||||||:::   |||:::      :::|||||
Db        224 uProThrSerSerValLysLeuLeuLeuGlnGlnAspThrAspLeuAlaHisLysAspIl 244

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
                |:::|||   |||||||||   ||||||:::      |||              |||
Db        244 eTyrGlyPheThrAlaGluGluTyrAlaSerPheAsnGlyPheThrMetTyrHis----- 262

Qy        949 ACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACC 994
                      ||||||            :::|||   :::   |||
Db        263 ----------HisIleThrAlaAsnAsnGluAsnLysLysLysThr 274
```

Search completed: July 15, 2004, 09:27:06
Job time : 242.464 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 08:45:37 ; Search time 31.051 Seconds
(without alignments)
12318.669 Million cell updates/sec

Title: US-09-602-362E-26
Perfect score: 6448
Sequence: 1 caagagcttggcgatacaga.....tgacttctaaattgaaggaa 3673

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093625_21988/app_query.fasta_1.10325
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09602362_@CGN_1_1_134_@runat_15072004_093625_21988 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	606	9.4	992	1	AN18_HUMAN	Q8ivf6 homo sapien
2	588	9.1	584	1	POTE_HUMAN	Q86yr6 homo sapien
3	362	5.6	201	1	ANR7_HUMAN	Q92527 homo sapien
4	347.5	5.4	1427	1	REST_HUMAN	P30622 homo sapien
5	339.5	5.3	1790	1	USO1_YEAST	P25386 saccharomyc
6	330	5.1	2663	1	CENE_HUMAN	Q02224 homo sapien
7	324.5	5.0	1972	1	MYHB_MOUSE	O08638 mus musculu
8	319	4.9	1960	1	MYH9_HUMAN	P35579 homo sapien
9	319	4.9	1972	1	MYHB_HUMAN	P35749 homo sapien
10	318.5	4.9	2116	1	MYS2_DICDI	P08799 dictyosteli
11	315	4.9	1972	1	MYHB_RABIT	P35748 oryctolagus
12	311	4.8	3259	1	GOB1_HUMAN	Q14789 homo sapien
13	310	4.8	2230	1	GOA4_HUMAN	Q13439 homo sapien
14	308.5	4.8	1961	1	MYH9_RAT	Q62812 rattus norv
15	300	4.7	1978	1	MYHB_CHICK	P10587 gallus gall
16	299.5	4.6	2871	1	DESP_HUMAN	P15924 homo sapien
17	297.5	4.6	1690	1	C190_DROME	Q9vje5 drosophila
18	297	4.6	2017	1	MYSN_DROME	Q99323 drosophila
19	296	4.6	1938	1	MYHD_HUMAN	Q9ukx3 homo sapien
20	293.5	4.6	1957	1	SPOF_SCHPO	Q10411 schizosacch
21	293	4.5	6632	1	UN89_CAEEL	O01761 caenorhabdi
22	290	4.5	1433	1	REST_CHICK	O42184 gallus gall
23	289	4.5	1875	1	MLP1_YEAST	Q02455 saccharomyc
24	289	4.5	3210	1	CENF_HUMAN	P49454 homo sapien
25	286	4.4	993	1	SCP1_MOUSE	Q62209 mus musculu
26	286	4.4	1364	1	KTN1_CHICK	Q90631 gallus gall
27	284.5	4.4	1727	1	ALM1_SCHPO	Q9utk5 schizosacch
28	284	4.4	3911	1	AKA9_HUMAN	Q99996 h a-kinase
29	283.5	4.4	1411	1	EEA1_HUMAN	Q15075 homo sapien
30	283	4.4	1357	1	KTN1_HUMAN	Q86up2 homo sapien
31	281	4.4	1391	1	MST2_DROHY	Q08696 drosophila
32	279	4.3	1959	1	MYH9_CHICK	P14105 gallus gall
33	278.5	4.3	1330	1	KTN1_VULVU	O97961 vulpes vulp
34	275	4.3	1583	1	GCC2_HUMAN	Q8iwj2 homo sapien
35	273	4.2	3924	1	ANK2_HUMAN	Q01484 homo sapien
36	271	4.2	976	1	SCP1_HUMAN	Q15431 homo sapien
37	270	4.2	1756	1	PEPL_HUMAN	O60437 homo sapien
38	268.5	4.2	1453	1	Y373_BOVIN	Q9tu23 bos taurus
39	268.5	4.2	1928	1	MYS1_YEAST	P08964 saccharomyc
40	268	4.2	1679	1	GCC2_MOUSE	Q8chg3 mus musculu
41	266.5	4.1	1679	1	YIO9_YEAST	P40457 saccharomyc
42	265.5	4.1	2442	1	CEP2_HUMAN	Q9bv73 homo sapien
43	264.5	4.1	1976	1	MYHA_RAT	Q9jlt0 rattus norv
44	264.5	4.1	2238	1	GOA4_MOUSE	Q91vw5 mus musculu
45	264.5	4.1	2349	1	TPR_HUMAN	P12270 homo sapien

ALIGNMENTS

RESULT 1
AN18_HUMAN
ID AN18_HUMAN STANDARD; PRT; 992 AA.
AC Q8IVF6; Q7Z468;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ankyrin repeat domain protein 18A.
GN ANKRD18A OR KIAA2015.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 1-349 FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";

```
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB095935; BAC23111.1; ALT_INIT.
DR   EMBL; BC056266; AAH56266.1; ALT_TERM.
DR   Genew; HGNC:23643; ANKRD18A.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 6.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       67     96       ANK 1.
FT   REPEAT      100    129       ANK 2.
FT   REPEAT      133    162       ANK 3.
FT   REPEAT      166    195       ANK 4.
FT   REPEAT      199    228       ANK 5.
FT   CONFLICT    130    130       K -> E (IN REF. 2).
FT   CONFLICT    288    288       MISSING (IN REF. 2).
SQ   SEQUENCE   992 AA;  115664 MW;  B9B288F087340D9A CRC64;

Alignment Scores:
Pred. No.:              6.48e-23        Length:         992
Score:                  606.00          Matches:        291
Percent Similarity:     40.86%          Conservative:   203
Best Local Similarity:  24.07%          Mismatches:     391
Query Match:            9.40%           Indels:         325
DB:                     1               Gaps:           53

US-09-602-362E-26 (1-3673) x AN18_HUMAN (1-992)

Qy        190 CTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCCTTCAGCGAA 249
Db         15 LeuLeuSerSerMetAspGlnGluTyrAlaGlyProGly---------------------- 27

Qy        250 CGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTCGGGGATCTAGGGAAGATCCAT 309
Db         28 ------TyrAspIleArgAspTrp-----------------GluLeuArgLysIleHis 39

Qy        310 ACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAGATGACAGTAGGGAAGAAGCCC 369
Db         40 ArgAlaAlaIleLysGlyAspAlaAlaGluValGluArg---CysLeuThrArgArgPhe 58

Qy        370 GTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAAT 429
Db         59 ArgAspLeuAspAlaArgAspArgLysAspArgThrValLeuHisLeuAlaCysAlaHis 78

Qy        430 GGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGAT 489
Db         79 GlyArgValGlnValValThrLeuLeuLeuHisArgArgCysGlnIleAspIleCysAsp 98

Qy        490 GGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAA 549
Db         99 ArgLeuAsnArgThrProLeuMetLysAlaValHisSerGlnGluGluAla-CysAlaIl 118

Qy        550 TATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCT 609
Db        118 eValLeuLeuGluCysGlyAlaAsnProAsnIleLysAspIleTyrGlyAsnThrAlaLe 138

Qy        610 CCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGC 669
Db        138 uHisTyrAlaValTyrAsnLysGlyThrSerLeuAlaGluArgLeuLeuSerHisHisAl 158

Qy        670 AGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAG 729

Db        158 aAsnIleGluAlaLeuAsnLysGluGlyAsnThrProLeuLeuPheAlaIleAsnSerAr 178

Qy        730 AAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGA 789
Db        178 gArgGlnHisMetValGluPheLeuLeuLysAsnGlnAlaAsnIleHisAlaValAspAs 198

Qy        790 GTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCAT 849
Db        198 nPheLysArgThrAlaLeuIleLeuAlaValGlnHisAsnLeuSerSerIleValThrLe 218

Qy        850 GCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACG 909
Db        218 uLeuLeuGlnGlnAsnIleArgIleSerSerGlnAspMetPheGlyGlnThrAlaGluAs 238

Qy        910 TTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAA 969
Db        238 pTyrAlaLeuCysSerAspLeuArgSerIleArgGlnGlnIleLeuGluHisLysAsnLy 258

Qy        970 ATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGA 1029
Db        258 sMetLeuLysAsnHisLeuArgAsnAspAsnGlnGluThrAlaAlaMetLysPro-AlaA 278

Qy       1030 GGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACC 1089
Db        278 snLeu-----------LysLysArgLysGluArgAlaLysAla---------------- 288

Qy       1090 TGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGC 1149
Db        289 ------------------------------------------------------------G 289

Qy       1150 AACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTAC 1209
Db        289 luHisAsnLeuLysValAlaSerGluGluLys--------------------------- 299

Qy       1210 AAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATCACATG 1269
Db        300 --GlnGluArgLeuGlnArgSerGluAsnLysGlnProGlnAsp---SerGlnSerTyrG 318

Qy       1270 GGAGGAAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAATAAAA- 1328
Db        318 lyLysLysLys-----------------AspAlaMetTyrGlyAsnPheMetLeuLysL 332

Qy       1329 -----CTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAAC 1383
Db        332 ysAspIleAlaMetLeuLysGluGlu-----------LeuTyrAlaIleLysAsnAspS 348

Qy       1384 ATCTACAAAAGCAAGTACAAATGTG----GATGTGAGTTCTGTAGAGCCTATATTCAGTC 1439
Db        348 erLeuArgLysGluLysLysTyrIleGlnGluIleLysSerIleThrGluIleAsnAlaA 368

Qy       1440 TTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATC 1499
Db        368 snPhe-----------------------------------------GluLysSerValArgL 375

Qy       1500 TTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTA 1559
Db        375 euAsnGluLysMetIleThrLysThrValAlaArg------------------------ 386

Qy       1560 CATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAG 1619
Db        387 --TyrSerGlnGlnLeuAsnAspLeu------LysAlaGluAsnAlaArgLeuAsnSerG 404

Qy       1620 AATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTT 1679
Db        404 luLeuGluLysGluLysHisAsnLys---------------------------------- 412

Qy       1680 CTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATA 1739
Db        413 ------------------------------------------------GluArgLeuGluA 417

Qy       1740 GAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTN---GAAATGC 1796
```

```
Db        417 laGluValGluSerLeu---------HisSerSerLeuAlaThrAlaIleAsnGluTyrA 434

Qy       1797 AAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTC 1856

Db        434 snGluIleValGluArgLysAspLeuGluLeu------------ValLeuTrpArgAla- 449

Qy       1857 AGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGA 1916

Db        450 ----------------------------AspAspValSerArgHisGluLysMetGlyS 460

Qy       1917 GTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAA- 1975

Db        460 erAsnIleSerGlnLeuThrAspLysAsnGluLeuLeuThrGluGlnValHisLysAlaA 480

Qy       1976 -----GAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTC 2030

Db        480 rgValLysPheAsnThrLeuLysGlyLysLeuArgGluThr------ArgAspAlaLeuA 498

Qy       2031 TGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACA 2090

Db        498 rgGluLysThr------------LeuAlaLeuGlySerValGlnLeuAspLeuArgGlnA 514

Qy       2091 GAGAA------------ACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGC 2138

Db        514 laGlnHisArgIleLysGluMetLysGlnMetHisProAsnGlyGluAlaLysGluSerG 534

Qy       2139 CTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAA 2198

Db        534 lnSerIleGlyLysGlnAsnSerLeuGluGluArgIle---------------------- 546

Qy       2199 CACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAG 2258

Db        547 ----------------------------------------------------ArgGlnG 549

Qy       2259 TTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGA 2318

Db        549 lnGluLeuGluAsnLeuLeuLeuGlu------------------------------- 557

Qy       2319 TGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTT 2378

Db        558 ----------------ArgGlnLeuGluAspAlaArgLysGluGlyAspAsnLysGluI 572

Qy       2379 TCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAAT 2438

Db        572 leValIleAsnIleHisArgAsp------CysLeuGlu---------------------- 582

Qy       2439 TCGATACCTTAAGTGGAAAA------TTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGA 2492

Db        583 -----------AsnGlyLysGluAspLeuLeuGluGluArgAsnLysGlu---LeuMetL 598

Qy       2493 AGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAG 2552

Db        598 ysGluTyrAsnTyrLeuLysGluLysLeuLeuGlnCysGluLysGluLysAlaGluArgG 618

Qy       2553 AAACA------------TTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTC 2600

Db        618 luValIleValArgGluPheGlnGluGluLeuValAspHisLeuLys---ThrPheSerI 637

Qy       2601 TTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATC 2660

Db        637 leSerGluSerProLeuGluGlyThrSerHisCys--------------HisIleAsnL 652

Qy       2661 TTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCA 2720

Db        652 euAsnGlu-----------------------ThrTrpThrSerLysLysLysLeuPheG 664

Qy       2721 TTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAA 2780

Db        664 lnValGluIleGlnProGluGlu------------------------------LysH 673

Qy       2781 AAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATT 2840

Db        673 isGluGluPheArgLysLeu------PheGluLeuIleSerLeu---------LeuAsnT 688

Qy       2841 ATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAG 2900

Db        688 yrThrAla--------------------AspGlnIleArgLysLysAsnArgGluLeuG 701

Qy       2901 AAGAGTCTCCTGAAAAGCCTTCTCAC---------TTTGAGCCTGCCACTGAAATGCAAA 2951

Db        701 luGlu------GluAlaThrGlyTyrLysLysCysLeuGluMetThrIleAsnMetLeuA 719

Qy       2952 ACTCTGTTCCAAATAAAGGCTTAGAATGGAAG---------AATAAACAAACATTGAGAG 3002

Db        719 snAlaPheAlaAsnGluAspPheAsnCysHisGlyAspLeuAsnThrAspGlnLeuLysM 739

Qy       3003 CAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGG 3062

Db        739 etAsp---IleLeuPheLysLysLeuLysGlnLysPheAsnAspLeuValAlaGluLysG 758

Qy       3063 AACTTAAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGT 3122

Db        758 luAlaValSerSerGluCysValAsnLeuAlaLysAspAsnGlu---------------- 772

Qy       3123 TTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA---------TCACAGT 3173

Db        773 -----ValLeuHisGlnGluLeuLeuSerMetArgAsnValGlnGluLysCysGluLysL 791

Qy       3174 TAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA------------- 3220

Db        791 euGluLysAspLysLysMetLeuGluGluGluValLeuAsnLeuLysThrHisMetGluL 811

Qy       3221 --------TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGAT------ATAT 3266

Db        811 ysAspMetValGluLeuGlyLysLeuGlnGluTyrLysSerGluLeuAspGluArgAlaV 831

Qy       3267 TAAAAGAAAAAATTAGACCCGAAGAG---CAACTTAGGAAAAAGTTAGAAGTGAAACACC 3323

Db        831 alGlnGluIleGluLysLeuGluGluIleHisLeuGlnLysGlnAlaGluTyrGluLysG 851

Qy       3324 AACTTGAACAG--------------------ACTCTCAGAATACAAGATATAGAATTGA 3362

Db        851 lnLeuGluGlnLeuAsnLysAspAsnThrAlaSerLeuLysLysLysGluLeuThrLeuL 871

Qy       3363 AAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAA---------AGTGAAA 3413

Db        871 ysAspValGluCysLysPheSerLysMetLysThrAlaTyrGluGluValThrThrGluL 891

Qy       3414 ATGATCTCTTTCATGAA--------------------------AATTGCATGTTGAAA- 3445

Db        891 euGluGluPheLysGluValPheAlaAlaAlaValLysAlaAsnAsnSerMetSerLysL 911

Qy       3446 -----------------AAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAAC 3488

Db        911 ysLeuMetLysSerAspLysLysIleAlaValIleSerThrLysLeuPheThrGluLys- 930

Qy       3489 ATCAACACCAGGTGAAGGAAAATAAATACTTT--------------------------G 3521

Db        931 --------------GlnArgMetLysTyrPheLeuSerThrLeuProThrArgProGluP 946

Qy       3522 AGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAAC 3581

Db        946 roGluLeuProCysValGluAsnLeuAsn--------------SerIleGluLeuAsnA 961

Qy       3582 AGAAAACAGTAACAAAAAGGGCA 3604

Db        961 rgLysTyrIleProLysThrAla 968
```

RESULT 2
POTE_HUMAN
ID   POTE_HUMAN   STANDARD;   PRT;   584 AA.
AC   Q86YR6;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   POTE protein (Prostate, ovary, testis expressed protein).

```
GN   POTE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RC   TISSUE=Prostate;
RX   MEDLINE=22388248; PubMed=12475935;
RA   Bera T.K., Zimonjic D.B., Popescu N.C., Sathyanarayana B.K., Kumar V.,
RA   Lee B., Pastan I.;
RT   "POTE, a highly homologous gene family located on numerous
RT   chromosomes and expressed in prostate, ovary, testis, placenta, and
RT   prostate cancer.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16975-16980(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA   Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA   Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
CC   -!- SUBCELLULAR LOCATION: Basal and terminal prostate epithelial
CC       cells.
CC   -!- TISSUE SPECIFICITY: Expressed in prostate, ovary, testis, placenta
CC       and prostate cancer cell lines.
CC   -!- SIMILARITY: Contains 6 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY172978; AAO23914.1; -.
DR   EMBL; AP001465; -; NOT_ANNOTATED_CDS.
DR   MIM; 607549; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 7.
DR   SMART; SM00248; ANK; 6.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Repeat; ANK repeat; Polymorphism.
FT   REPEAT      172    201       ANK 1.
FT   REPEAT      205    234       ANK 2.
FT   REPEAT      238    267       ANK 3.
FT   REPEAT      271    300       ANK 4.
FT   REPEAT      304    333       ANK 5.
FT   REPEAT      337    366       ANK 6.
FT   VARIANT     113    113       S -> G (in dbSNP:6517869).
FT                                /FTId=VAR_016242.
FT   VARIANT     135    135       V -> I (in dbSNP:6517870).
FT                                /FTId=VAR_016243.
SQ   SEQUENCE   584 AA;  66378 MW;  CB0199694EC63340 CRC64;
```

Alignment Scores:			
Pred. No.:	4.87e-22	Length:	584
Score:	588.00	Matches:	176
Percent Similarity:	46.36%	Conservative:	98
Best Local Similarity:	29.78%	Mismatches:	208
Query Match:	9.12%	Indels:	109

```
DB:                    1              Gaps:          16

US-09-602-362E-26 (1-3673) x POTE_HUMAN (1-584)

Qy        124 GGGGCTGGGGAAGGGCGAGCGGGAGGCGCGGGCTCTCTCTAGCAGGGGGCTGCAGCCATG 183
              |||:::|||          |||  :::|||                      :::   |||
Db         74 GlySerGlyThrSerAsnValGlyThrSerGlyAsp------HisGluAsnSerPheMet 91

Qy        184 AAGAGGCTCTTAGCTGCCGCTGGCAAGGGC--------------------GTGCGGGGC 222
              |||   |||   :::      ||||||                        ||||||
Db         92 LysMetLeuArgSerLysMetGlyLysTrpCysCysHisCysProCysCysArgGly 111

Qy        223 CCGGAGCCCCCGAAC--------------CCCTTCAGCGAACGGGTCTACACTGAGAAG 267
                  |||                  :::                 :::   |||
Db        112 SerSerLysSerAsnValGlyAlaTrpGlyAspTyrAspHisSerAlaPheMetGluPro 131

Qy        268 GACTACGGGACCATCTACTTCGGGGATCTAGGGAAGATCCATACAGCTGCCTCCCGGGGC 327
              |||     :::          ||||||  |||:::|||  ||||||     |||
Db        132 ArgTyr---HisValArgArgGluAspLeuAspLysLeuHisArgAlaAlaTrpTrpGly 150

Qy        328 CAAGTCCAGAAGCTGGAGAAGATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGA 387
              :::|||   :::   :::   :::   ||      :::      ::::::||||||||
Db        151 LysValProArg---LysAspLeuIleValMetLeuArgAspThrAspMetAsnLysArg 169

Qy        388 GATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTA 447
              |||   :::|||||||||||||||||||  |||     ||||||::::::   ||||||
Db        170 AspLysGluLysArgThrAlaLeuHisLeuAlaSerAlaAsnGlyAsnSerGluValVal 189

Qy        448 ACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCT 507
                ||:::||||||:::|||  ||||||||||||||||   :::   ||||||
Db        190 GlnLeuLeuLeuAspArgArgCysGlnLeuAsnValLeuAspAsnLysLysArgThrAla 209

Qy        508 CTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGG 567
              |||:::||||||:::||||||:::   :::    |||       :::|||::::::   ||
Db        210 LeuIleLysAlaIleGlnCysGlnGluAspGlu-CysValLeuMetLeuLeuGluHisGl 229

Qy        568 TGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAG 627
              |||||||   |||       |||   |||||||||||||||||||||||||:::|||::
Db        229 yAlaAspArgAsnIleProAspGluTyrGlyAsnThrAlaLeuHisTyrAlaIleTyrAs 249

Qy        628 TGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAA 687
              :|||:::   ||||||          ||||||   |||||||||   ||||||   :::||
Db        249 nGluAspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyValAlaAspIleGluSerLysAs 269

Qy        688 CAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGA 747
              ||||      |||||||||||||||||   :::   ::::::   :::|||   |||::
Db        269 nLysCysGlyLeuThrProLeuLeuLeuGlyValHisGluGlnLysGlnGlnValValLy 289

Qy        748 ATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCT 807
              :||||||:::   |||   ||||||   |||       :::            ||||||||
Db        289 sPheLeuIleLysLysLysAlaAsnLeuAsnValLeuAspArgTyrGlyArgThrAlaLe 309

Qy        808 CATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGT 867
              |:::||||||:::|||   ||||||:::   ||||||  :::|||||:::||||||||
Db        309 uIleLeuAlaValCysCysGlySerAlaSerIleValAsnLeuLeuLeuGluGlnAsnVa 329

Qy        868 TGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGG 927
              |||||||  ::::::|||:::   |||   ||||||       ||||||   :::
Db        329 lAspValSerSerGlnAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerHi 349

Qy        928 AGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAA------TTACCTAAAAA 981
                :::   |||  :::   |||  ::::::      |||       |||       ::
Db        349 sHisHisValIleCysGluLeuLeuSerAspTyrLysGluLysGlnMetLeuLysIleSe 369

Qy        982 TCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTT 1041
              :   :::|||:::||||||||
Db        369 rSerGluAsnSerAsnProGlu---------------------------------------- 376

Qy       1042 GGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGC 1101
                                    :::   |||       :::   :::|||:::
Db        377 ------------------------GlnAspLeuLysLeuThrSerGluGluGluSerGl 388

Qy       1102 ACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAA 1161
```

```
                 ||||||
Db        388 nArgLeu---------------------------------------------------- 390

Qy       1162 GTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATC 1221
                 :::   |||          |||   |||:::
Db        391 ----LysValSerGluAsnSerGlnProGluLysMet---------------------- 401

Qy       1222 TGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGA 1281
                                                                     :::||
Db        402 ------------------------------------------------------SerGlnGl 404

Qy       1282 AACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGA 1341
                 |      :::      :::|||        |||               |||   :::::
Db        404 uProGluIleAsnLysAspCysAspArgGluVal-----------GluGluGluIleLy 420

Qy       1342 AAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTAC 1401
                 :|||      ||||||   :::      |||         |||:::     ::::::
Db        420 sLysHisGlySerAsnProValGlyLeuProGluAsnLeuThrAsnGlyAlaSerAlaGl 440

Qy       1402 AAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGA 1461
                 |||   |||              :::   :::             :::|||     ||
Db        440 yAsnGlyAsp-----------AspGlyLeuIleProGlnArgArgSerArgLysProGl 456

Qy       1462 AAATTCACAGTGTACAAAAGTTGAG-----------------GAAGACTTTAATCTTGC 1503
                 ||||   |||          |||                 ::::::   |||
Db        456 uAsnGlnGlnPheProAspThrGluAsnGluGluTyrHisSerAspGluGlnAsnAspTh 476

Qy       1504 TACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGAT-------- 1555
                 |||   :::|||:::         ||||||              |||
Db        476 rArgLysGlnLeuSerGluGlu------GlnAsnThrGlyIleSerGlnAspGluIleLe 494

Qy       1556 -GCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCC 1614
                  ||||||   |||:::            |||:::      ::::::
Db        494 uThrAsnLysGlnLysGlnIleGluValAlaGluGlnLysMetAsnSerGluLeuSerLe 514

Qy       1615 ATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGA 1674
                 |||      |||        :::|||:::          ::::::          |||   ||
Db        514 uSerHisLysLys---------GluGluAspLeuLeuArgGluAsnSerValLeuGlnGl 531

Qy       1675 GAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGAT 1734
                 |      |||      ::::::   :::   |||:::   :::|||      :::
Db        531 uGluIleAlaValLeuArgLeuGluLeuAspGluThrLysHisGlnAsnGlnLeuArgGl 551

Qy       1735 AAATAGA---------GAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGC 1785
                 |||:::         |||:::|||   :::   ||||||                  ||
Db        551 uAsnLysIleLeuGluGluIleGluSerValLysGluLysThrAspLysLeuLeuArgAl 571

Qy       1786 CGTNGAAATGCAAAAGACTGTTCCAAATAAA 1816
                 |:::::::::   :::            |||
Db        571 aMetGlnLeuAsnGluGluAlaLeuThrLys 581
```

RESULT 3
ANR7_HUMAN
ID   ANR7_HUMAN   STANDARD;   PRT;   201 AA.
AC   Q92527; Q96QN1; Q9UDM3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat domain protein 7 (Testis-specific protein TSA806).
GN   ANKRD7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, and Testis;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE OF 51-201 FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=96411689; PubMed=8812458;
RA   Ozaki K., Kuroki T., Hayashi S., Nakamura Y.;
RT   "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
RT   a differential mRNA display method.";
RL   Genomics 36:316-319(1996).
RN   [3]
RP   SEQUENCE OF 8-138 FROM N.A.
RA   Courtney L., Wohldmann P., Johnson D., Tucci S.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- TISSUE SPECIFICITY: Testis specific.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; BC032799; AAH32799.1; -.
DR   EMBL; BC007020; AAH07020.1; -.
DR   EMBL; D78334; BAA11348.1; -.
DR   EMBL; AC007874; AAF19231.1; -.
DR   Genew; HGNC:18588; ANKRD7.
DR   HSSP; P25963; 1IKN.
DR   GO; GO:0008584; P:male gonad development; TAS.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT   5   34   ANK 1.
FT   REPEAT   38   67   ANK 2.
FT   REPEAT   71   100   ANK 3.
FT   REPEAT   104   133   ANK 4.
FT   REPEAT   137   166   ANK 5.
FT   CONFLICT   51   55   NEDCA -> MRIVP (IN REF. 2).
SQ   SEQUENCE   201 AA; 22588 MW; 30C241795D2C049F CRC64;

Alignment Scores:

Pred. No.:	5.98e-11	Length:	201
Score:	362.00	Matches:	77
Percent Similarity:	65.68%	Conservative:	34
Best Local Similarity:	45.56%	Mismatches:	57
Query Match:	5.61%	Indels:	1
DB:	1	Gaps:	0

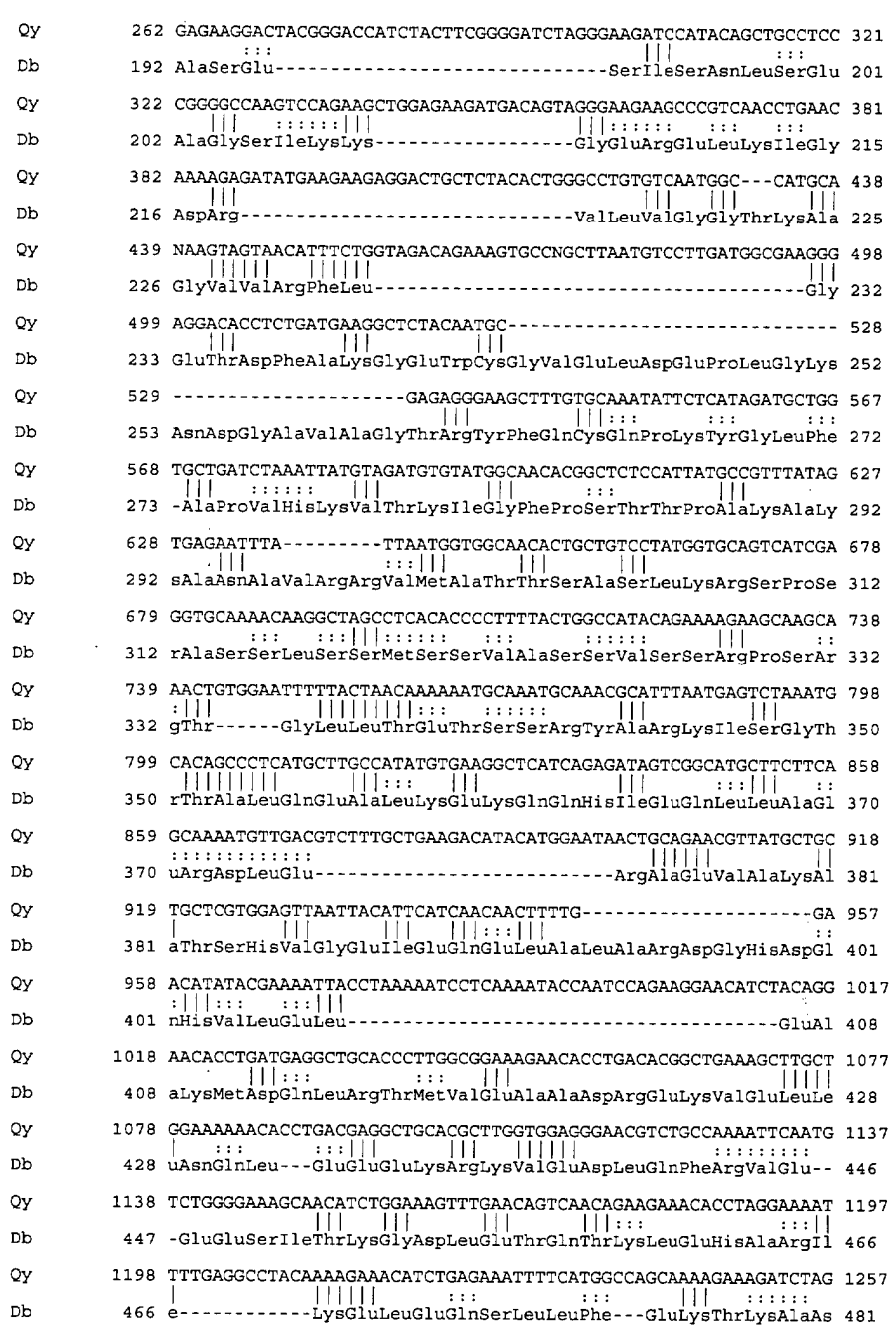
US-09-602-362E-26 (1-3673) x ANR7_HUMAN (1-201)

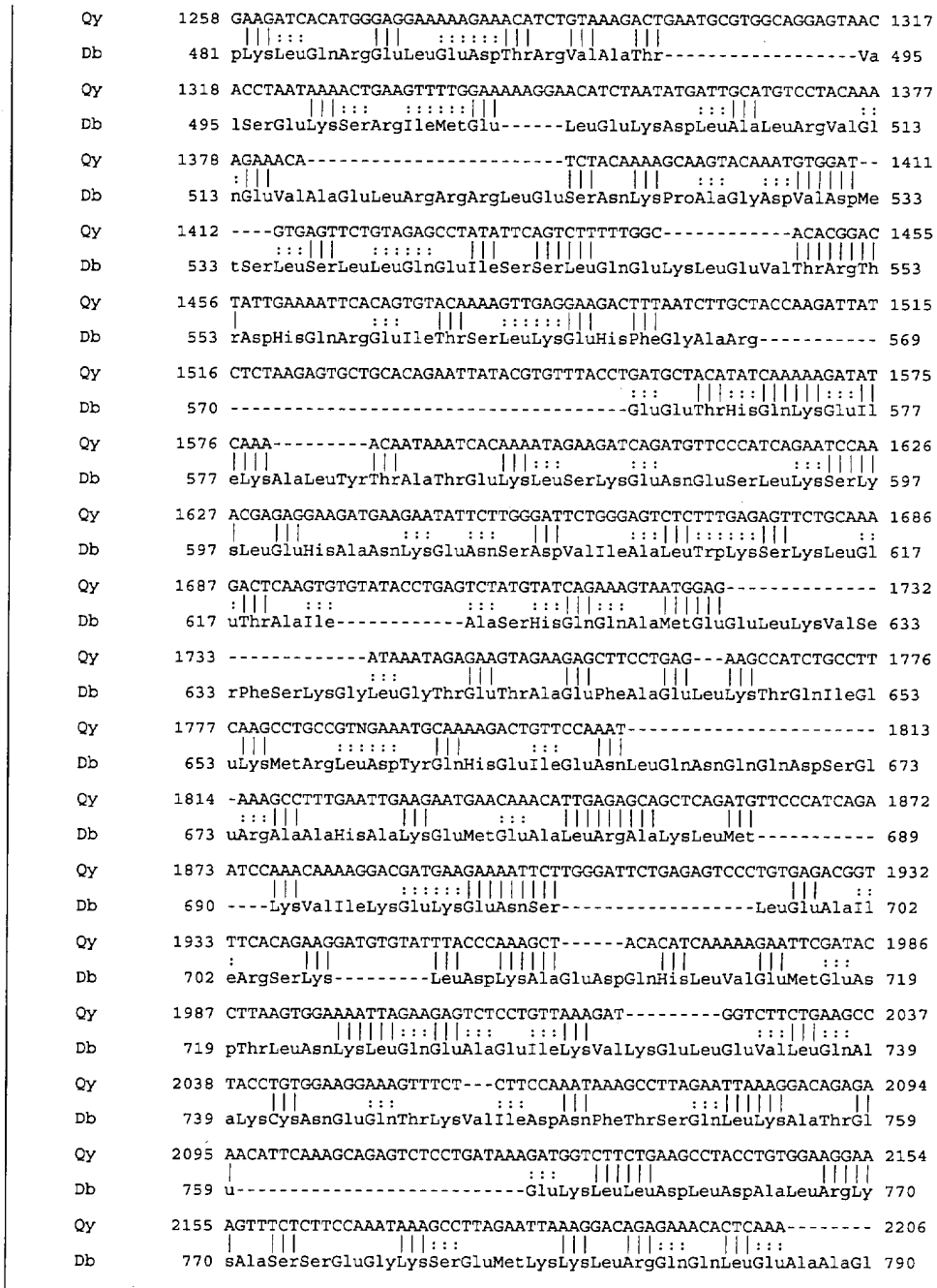
```
Qy        385 AGAGATATGAAGAAGAGGACTGCTCTACACTGGGCCTGTGTCAATGGCCATGCANAAGTA 444
                 :::|||   |||  ||||||  ||||||  ||||||  |||||||||      |||
Db          2 GlnAspLysLysTyrArgThrProLeuHisLeuAlaCysAlaAsnGlyHisThrAspVal 21
```

```
Qy        445 GTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACA 504
Db         22 ValLeuPheLeuIleGluGlnGlnCysLysIleAsnValArgAspSerGluAsnLysSer 41

Qy        505 CCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGC 564
Db         42 ProLeuIleLysAlaValGlnCysGlnAsnGluAsp-CysAlaThrIleLeuLeuAsnPh 61

Qy        565 TGGTGCTGATCTAAATTATGTAGATGTGTATGGCAACACGGCTCTCCATTATGCCGTTTA 624
Db         61 eGlyAlaAspProAspLeuArgAspIleArgTyrAsnThrValLeuHisTyrAlaValCy 81

Qy        625 TAGTGAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCA 684
Db         81 sGlyGlnSerLeuSerLeuValGluLysLeuLeuGluTyrGluAlaAspLeuGluAlaLy 101

Qy        685 AAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGT 744
Db        101 sAsnLysAspGlyTyrThrProLeuLeuValAlaValIleAsnAsnAsnProLysMetVa 121

Qy        745 GGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGC 804
Db        121 lLysPheLeuLeuGluLysGlyAlaAspValAsnAlaSerAspAsnTyrGlnArgThrAl 141

Qy        805 CCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAA 864
Db        141 aLeuIleLeuAlaValSerGlyGluProProCysLeuValLysLeuLeuLeuGlnGlnGl 161

Qy        865 TGTTGACGTCTTTGCTGAAGACATA 889
Db        161 yValGluLeuCysTyrGluGlyIle 169
```

```
RESULT 4
REST_HUMAN
ID   REST_HUMAN     STANDARD;      PRT;  1427 AA.
AC   P30622;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Restin (Cytoplasmic linker protein-170 alpha-2) (CLIP-170) (Reed-
DE   Sternberg intermediate filament associated protein).
GN   RSN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood monocytes;
RX   MEDLINE=92289675; PubMed=1600942;
RA   Bilbe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA   Cerletti N., Sorg C., Odink K., Tarcsay L., Wiesendanger W.,
RA   de Wolf-Peeters C., Shipman R.;
RT   "Restin: a novel intermediate filament-associated protein highly
RT   expressed in the Reed-Sternberg cells of Hodgkin's disease.";
RL   EMBO J. 11:2103-2113(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92405160; PubMed=1356075;
RA   Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
RT   "CLIP-170 links endocytic vesicles to microtubules.";
RL   Cell 70:887-900(1992).
CC   -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
CC       THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
CC       CYTOSKELETON.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P30622-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P30622-2; Sequence=VSP_000765;
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
CC       OF HODGKIN'S DISEASE.
CC   -!- SIMILARITY: Contains 2 CAP-Gly domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X64838; CAA46050.1; -.
DR   EMBL; M97501; AAA35693.1; -.
DR   PIR; S22695; S22695.
DR   Genew; HGNC:10461; RSN.
DR   MIM; 179838; -.
DR   GO; GO:0005768; C:endosome; TAS.
DR   GO; GO:0005882; C:intermediate filament; TAS.
DR   GO; GO:0015630; C:microtubule cytoskeleton; TAS.
DR   GO; GO:0008017; F:microtubule binding; TAS.
DR   GO; GO:0006899; P:nonselective vesicle transport; TAS.
DR   InterPro; IPR000938; CAP-Gly.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF01302; CAP_GLY; 2.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS00845; CAP_GLY_1; 2.
DR   PROSITE; PS50245; CAP_GLY_2; 2.
KW   Cytoskeleton; Microtubule; Coiled coil; Repeat; Alternative splicing.
FT   DOMAIN       78    120       CAP-GLY 1.
FT   DOMAIN      143    204       SER-RICH.
FT   DOMAIN      232    274       CAP-GLY 2.
FT   DOMAIN      304    331       SER-RICH.
FT   DOMAIN      350   1342       COILED COIL (POTENTIAL).
FT   DOMAIN     1408   1421       CCHC-BOX.
FT   VARSPLIC    457    491       Missing (in isoform Short).
FT                                /FTId=VSP_000765.
FT   CONFLICT   1069   1069       D -> E (IN REF. 2).
SQ   SEQUENCE   1427 AA;  160989 MW;  0A4F166DD94254E8 CRC64;
```

```
Alignment Scores:
Pred. No.:                3.32e-10        Length:          1427
Score:                    347.50          Matches:         304
Percent Similarity:       36.79%          Conservative:    207
Best Local Similarity:    21.89%          Mismatches:      510
Query Match:              5.39%           Indels:          368
DB:                       1               Gaps:            67

US-09-602-362E-26 (1-3673) x REST_HUMAN (1-1427)

Qy          1 CAAGAGCTTGGCGATACAGAAATTTCTGCTGGTGTTGGGGCGGGT---------- 45
Db         75 GlnPheLeuGlyGluThrGlnPheAlaProGlyGlnTrpAlaGlyIleValLeuAspGlu 94

Qy         46 ---GCGGGAACTGAAGACGGGCGAGTGCGAGCCGGG---------- 78
Db         95 ProIleGlyLysAsnAspGlySerVal---AlaGlyValArgTyrPheGlnCysGluPro 113

Qy         79 ---GGCGGGTGCTGGGGAAGGGTAAGCGGGAAGCGAGGGCGAGGGGTAGGGGCTGGGGAA 135
Db        114 LeuLysGlyIlePheThrArgProSer---LysLeuThrArgLysValGlnAla---Glu 131

Qy        136 GGGCGAGCGGGAGGCGCGGGCTCTCTCTAGCAGGGGGCTGCAGCCATGAAGAGGCTCTTA 195
Db        132 AspGluAlaAsnGlyLeuGlnThrThrProAlaSerArgAlaThrSerProLeuCysThr 151

Qy        196 GCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAAC---------- 237
Db        152 SerThrAlaSerMetValSerSerSerProSerThrProSerAsnIleProGlnLysPro 171

Qy        238 ----------CCCTTCAGCGAACGGGTCTACACT 261
Db        172 SerGlnProAlaAlaLysGluProSerAlaThrProProIleSerAsnLeuThrLysThr 191
```



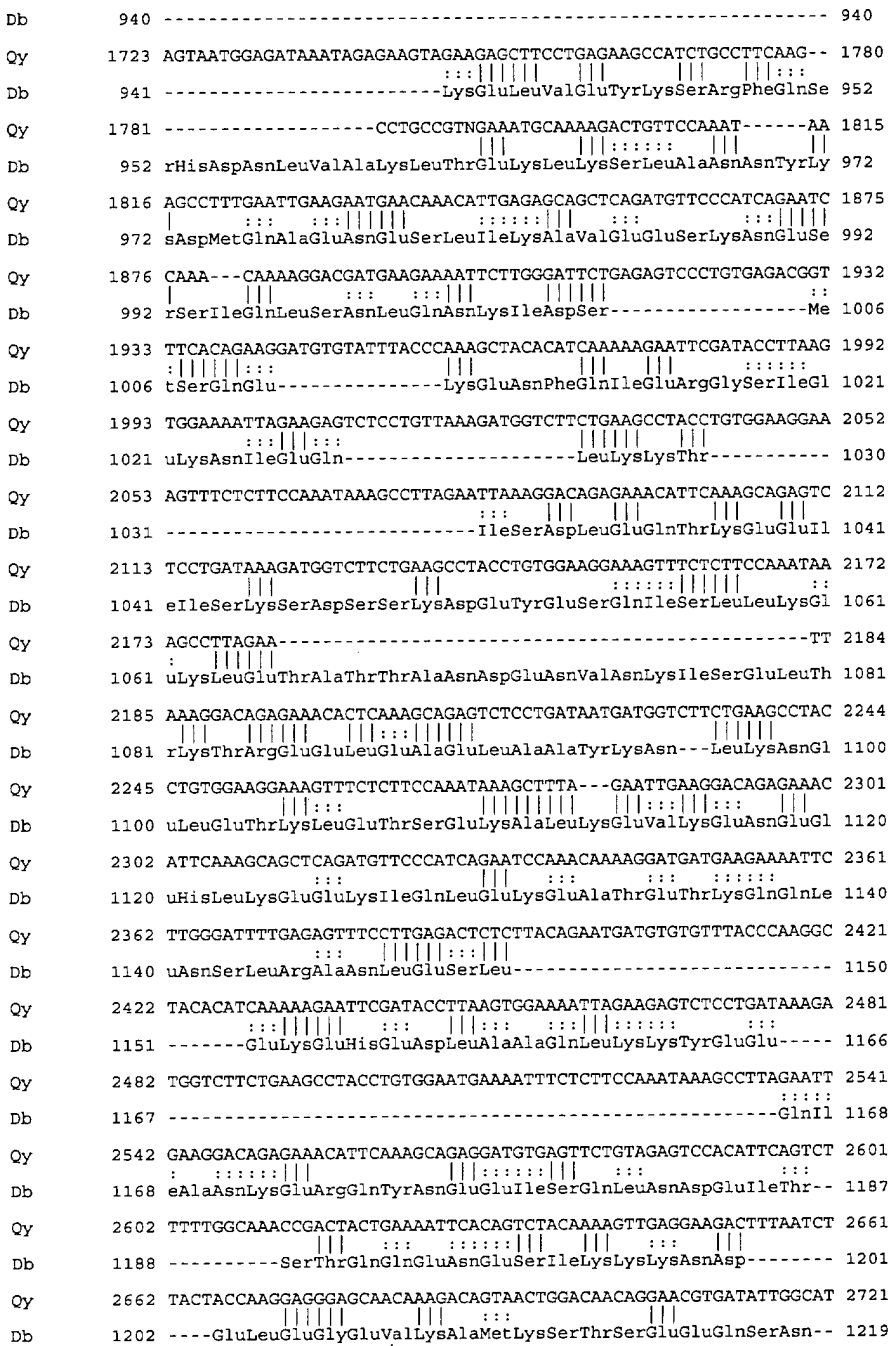


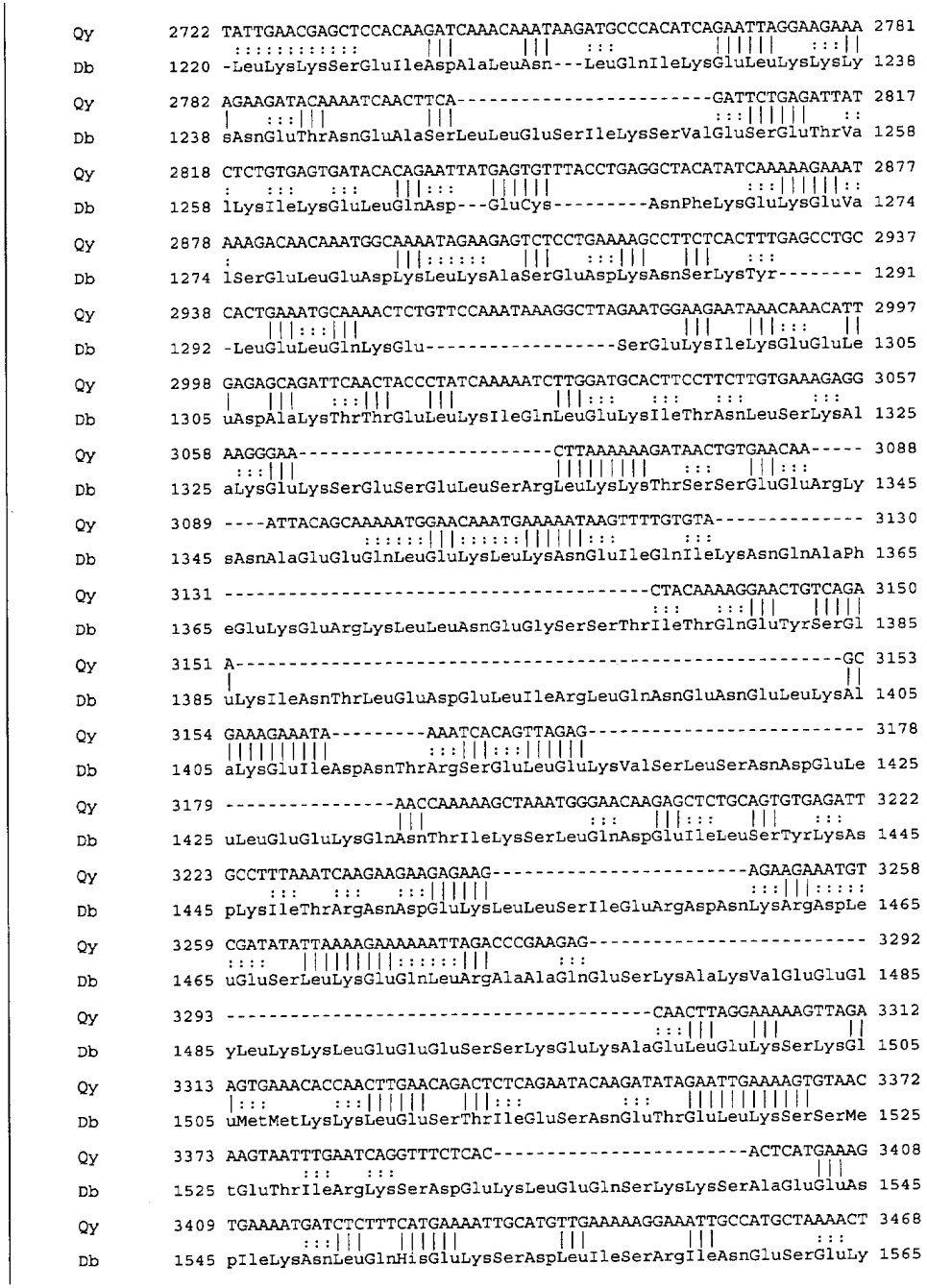
```
                                                 ||||||||                 ::
Db        790 uLysGlnIleLysHisLeuGluIleGluLysAsnAlaGluSerSerLysAlaSerSerIl 810

Qy       2233 TCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGA 2292
              :   :::         ||||||::::::   |||   |||              |||   ::
Db        810 eThrArgGluLeuGlnGlyArgGluLeuLysLeuThrAsnLeuGlnGluAsnLeuSerGl 830

Qy       2293 C------------AGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACA 2340
              :            :::||||||   :::                          |||
Db        830 uValSerGlnValLysGluThrLeuGluLysGluLeuGlnIleLeuLysGluLysPheAl 850

Qy       2341 AAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAA 2400
               :::      ||||||                 |||      ||||||:::
Db        850 aGluAlaSerGluGluAlaValSerValGlnArgSerMetGlnGluThrVal-------- 867

Qy       2401 TGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAA------TTCGATACCTTAAGTGG 2454
                                        ||||||||||||      |||:::   ||||||
Db        868 ----------------AsnLysLeuHisGlnLysGluGluGlnPheAsnMetLeuSerSe 882

Qy       2455 AAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAAT 2514
              ||||||:::                    |||   :::              |||:::
Db        882 rAspLeuGluLys-----------------LeuArgGluAsnLeuAlaAspMetGluAl 896

Qy       2515 TTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGT 2574
                        |||      :::   ::::::   :::      ||||||::::::
Db        896 aLysPheArgGluLys------AspGluArgGluGluGlnLeuIleLysAlaLysGlu-- 913

Qy       2575 GAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAAT---TCACA 2631
                    :::|||:::      :::          |||   :::   :::|||   |||||
Db        914 ----LysLeuGluAsnAspIleAlaGluIleMetLysMetSerGlyAspAsnSerSerGl 932

Qy       2632 GTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGT 2691
              |   ||||||:::   ::::::      |||
Db        932 nLeuThrLysMetAsnAspGluLeuArgLeu---------------------------- 942

Qy       2692 AACTGGACAACAGGAACGTGATATT---------------------------------- 2716
                        :::|||||||||:::
Db        943 ----------LysGluArgAspValGluGluLeuGlnLeuLysLeuThrLysAlaAsnGl 959

Qy       2717 -------GGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA-- 2767
                     ::::::::::::   :::|||   |||   |||          |||
Db        959 uAsnAlaSerPheLeuGlnLysSerIleGluAspMetThrValLysAlaGluGlnSerGl 979

Qy       2768 ----------------------GAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGA 2805
                                    |||      :::||||||      :::      |||||
Db        979 nGlnGluAlaAlaLysLysHisGluGluGluLysLysGluLeuGluArgLysLeuSerAs 999

Qy       2806 TTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCT---GAGGCTAC 2862
              |   |||                :::|||   :::   :::|||          :::|||
Db        999 pLeuGlu------LysLysMetGluThrSerHisAsnGlnCysGlnGluLeuLysAlaAr 1017

Qy       2863 ATATCAAAAA---------GAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGA 2913
              |||::::::          |||   ||||||   :::   :::|||   ::::::   ::
Db       1017 gTyrGluArgAlaThrSerGluThrLysThrLysHisGluGluIleLeuGlnAsnLeuGl 1037

Qy       2914 AAAG--------------CCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958
              :|||                        :::   |||   |||   :::||||||
Db       1037 nLysThrLeuLeuAspThrGluAspLysLeuLysGlyAlaArgGlu---GluAsnSerGl 1056

Qy       2959 TCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCT 3018
                 :::   ||||||          ||||||      :::|||   :::
Db       1056 yLeuLeuGlnGluLeuGluGluLeuArgLysGlnAlaAspLysAlaLysAlaAlaGlnTh 1076

Qy       3019 ATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAA 3078
              ::::::      ||||||:::          |||          ::::::   |||:::
Db       1076 rAlaGlu------AspAlaMetGlnIleMetGlu---------GlnMetThrLysGluLy 1091

Qy       3079 CTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAA 3138
                 |||   :::   |||   :::|||      |||                  ||||||
Db       1091 sThrGluThrLeu---AlaSerLeuGluAspThrLysGlnThrAsnAlaLysLeuGlnAs 1110

Qy       3139 GGAACTGTCAGAAGCGAAAGAA------ATAAAATCACAGTTAGAGAACCAAAAAGCTAA 3192
                 ||||||          ||||||      :::|||:::      |||      |||:::||
```

```
Db       1110 nGluLeuAspThrLeuLysGluAsnAsnLeuLysAsnValGluGluLeuAsnLysSerLy 1130

Qy       3193 ATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAA---GAAGAGAAGAG 3249
              |          ||||||   :::|||              ||||||:::   ||||||   ||
Db       1130 s---------GluLeuLeuThrValGlu---------AsnGlnLysMetGluGluPheAr 1144

Qy       3250 AAGAAATGTCGATATATTAAAA---------------------------------------- 3271
              |:::   ::::::   ||||||
Db       1144 gLysGluIleGluThrLeuLysGlnAlaAlaAlaGlnLysSerGlnGlnLeuSerAlaLe 1164

Qy       3272 -------GAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAG----------------- 3307
                     |||   ::::::      |||:::|||   :::
Db       1164 uGlnGluGluAsnValLysLeuAlaGluGluLeuGlyArgSerArgAspGluValThrSe 1184

Qy       3308 ---------------------------------------------TTAGAAGTGAAACA 3321
                                                           ||||||:::|||
Db       1184 rHisGlnLysLeuGluGluGluArgSerValLeuAsnAsnGlnLeuGluMetLysLy 1204

Qy       3322 CCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTT 3381
               :::      :::   ::::::      |||   |||      |||:::      :::::
Db       1204 sArgGluSerLysPheIleLysAspAlaAspGluGluLysAlaSerLeuGlnLysSerIl 1224

Qy       3382 GAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTT 3441
              ::::   ::::::            :::|||   |||          |||         ||
Db       1224 eSer---IleThrSerAlaLeuLeuThrGluLysAspAlaGluLeuGluLys------Le 1241

Qy       3442 GAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGT 3501
              |:::   |||:::   :::|||:::   |||   |||:::   |||      |||   ||
Db       1241 uArgAsnGluValThrValLeuArgGlyGluAsnAlaSerAlaLysSerLeuHisServa 1261

Qy       3502 GAAGGAAAATAAATACTTTGAGGACATTAAGATT---TTACAAGAAAAGAATGCTGAACT 3558
              |   :::            :::   :::|||:::   |||:::   ||||||   |||||
Db       1261 lValGlnThrLeuGluSerAspLysValLysLeuGluLeuLysValLysAsnLeuGluLe 1281

Qy       3559 TCAAATGACCCTAAAACTGAAACAG 3583
              |||:::            :::|||
Db       1281 uGlnLeuLysGluAsnLysArgGln 1289
```

```
RESULT 5
USO1_YEAST
ID   USO1_YEAST     STANDARD;      PRT;  1790 AA.
AC   P25386;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Intracellular protein transport protein USO1.
GN   USO1 OR INT1 OR YDL058W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=91185402; PubMed=2010462;
RA   Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA   Yamasaki M.;
RT   "A cytoskeleton-related gene, uso1, is required for intracellular
RT   protein transport in Saccharomyces cerevisiae.";
RL   J. Cell Biol. 113:245-260(1991).
RN   [2]
RP   SEQUENCE OF 782-1790 FROM N.A.
RA   Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA   Kendrick K.E.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-8 FROM N.A.
RA   Bai Y., Symington L.S.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Required for protein transport from the ER to the Golgi
CC       complex.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC       MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
```

```
CC       ER AND THE GOLGI COMPLEX.
CC   -!- DOMAIN: The rodlike tail sequence is highly repetitive, composed
CC       of an heptapeptide repeat pattern characteristic of alpha-helical
CC       coiled coils. May form filamentous structures in the cell.
CC   -!- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54378; CAA38253.1; -.
DR   EMBL; L03188; AAB00143.1; -.
DR   EMBL; U53668; AAB66659.1; -.
DR   GermOnline; 140300; -.
DR   SGD; S0002216; USO1.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR002017; Spectrin.
DR   InterPro; IPR006955; Uso1_p115_C.
DR   InterPro; IPR006953; Uso1_p115_head.
DR   Pfam; PF04871; Uso1_p115_C; 1.
DR   Pfam; PF04869; Uso1_p115_head; 1.
KW   Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT   DOMAIN        1    724       GLOBULAR HEAD.
FT   DOMAIN      725   1790       COILED COIL (POTENTIAL).
FT   DOMAIN      465    487       CHARGED (HYPER-HYDROPHILIC).
FT   DOMAIN      991   1790       DISPENSABLE FOR THE PROTEIN FUNCTION.
FT   DOMAIN     1172   1786       ASP/GLU-RICH (ACIDIC).
FT   CONFLICT    847    847       G -> E (IN REF. 2).
FT   CONFLICT    924    924       E -> K (IN REF. 2).
FT   CONFLICT   1253   1253       V -> I (IN REF. 2).
FT   CONFLICT   1319   1319       I -> V (IN REF. 2).
FT   CONFLICT   1461   1461       N -> S (IN REF. 2).
FT   CONFLICT   1581   1581       G -> S (IN REF. 2).
FT   CONFLICT   1600   1600       I -> V (IN REF. 2).
FT   CONFLICT   1661   1661       R -> S (IN REF. 2).
FT   CONFLICT   1772   1772       D -> DEEDDEE (IN REF. 2).
SQ   SEQUENCE   1790 AA;  206424 MW;  6CE2B216E9FD4818 CRC64;

Alignment Scores:
Pred. No.:                 8.27e-10        Length:          1790
Score:                     339.50          Matches:         268
Percent Similarity:        36.07%          Conservative:    193
Best Local Similarity:     20.97%          Mismatches:      441
Query Match:               5.27%           Indels:          376
DB:                        1               Gaps:            58

US-09-602-362E-26 (1-3673) x USO1_YEAST (1-1790)

Qy        463 AGAAAGTGCCNGCTTAATGTCCTTGATGGC---------GAAGGGAGGACACCTCTGATG 513
              |||:::          ||||||       |||          |||          ||||||
Db        534 ArgGluIleThrArgAsnValThrThrGlyAsnAspLeuGluAspGluGluProLeu--- 552

Qy        514 AAGGCTCTACAATGCGAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGA 573
              ||||||:::|||        ||| |||          |||          |||||
Db        553 LysAlaIleGlnThrIleSerGlu-LeuLeuThrThrSerLeu------ThrAlaAlaAs 570

Qy        574 TCTAAATTATGTAGATGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAG-- 631
              |:::      :::   :::          |||   |||   |||   ::::::   :::
Db        570 pIleArg---IleProIleSerTyrLeuThrPheLeuIleTyrTrpLeuPheGlyAspPh 589

Qy        632 ------------------------AATTTATTAATGGTGGCAACACTGCTGTCCTATGG 666
                                      :::      ::::::   :::||||||||||:::
Db        589 eLysAlaThrAsnAspPheLeuSerAspLysSerValIleLysSerLeuLeuSerPheSe 609

Qy        667 TGCAGTCATC---------GAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTA---CT 714
                  |||          :::|||       |||       :::|||   ||||||   ::
Db        609 rTyrGlnIleGlnAspGluAspValThrIleLysCysLeuValThrMetLeuLeuGlyVa 629

Qy        715 GGCCATACAGAAAAGAAGCAAGCAAACT--------------------GTGGAATTTTT 753
              :|||   :::      ||||||::::::                              ||||||::
Db        629 lAlaTyrGluPheSerSerLysGluSerProPheProArgLysGluTyrPheGluPheIl 649

Qy        754 A---------CTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGC 804
              :          |||   |||:::       |||:::          :::   |||      ::
Db        649 eThrLysThrLeuGlyLysAspAsnTyrAlaSerArgIleLysGlnPheLysLysAspSe 669

Qy        805 CCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAA 864
              :          :::                 |||:::          |||   :::
Db        669 rTyrPheSerLysValAspMetAsnGluAspSerIleLeuThrProGluLeuAspGluTh 689

Qy        865 T---------------------------------GTTGACGTCTTTGCTGAAGACATACA 891
                                                :::   :::|||   |||:::|||::
Db        689 rGlyLeuProLysValTyrPheSerThrTyrPheIleGlnLeuPheAsnGluAsnIleTy 709

Qy        892 TGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTAC------ATTCATCA 945
              :   |||          |||   |||   :::          :::          |||   ::
Db        709 rArgIle---------ArgThrAlaLeuSerHisAspProAspGluGluProIleSerLy 726

Qy        946 ACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGG 1005
              :          |||   :::   ||||||   :::       |||   |||       :::||
Db        726 sIleSerPheGluGluValGluLysLeuGlnArg------GlnCysThrLysLeuLysGl 744

Qy       1006 AACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGC 1065
              |              |||          |||          |||          |||
Db        744 y---------------------GluIleThrSerLeuGlnThrGluThrGluSerThrHi 757

Qy       1066 TGAAAGCTTGCTGGAAAAA---------ACACCTGACGAGGCTGCACGCTTGGTGGAGGG 1116
               |||:::|||   ||||||          :::|||          |||   |||
Db        757 sGluAsnLeuThrGluLysLeuIleAlaLeuThrAsnGluHisLysGluLeuAspGlu-- 776

Qy       1117 AACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAG---TC 1173
                    |||   |||   |||          :::   |||          |||       ||
Db        777 ----------LysTyrGlnIleLeuAsnSerSerHisSerSerLeuLysGluAsnPheSe 793

Qy       1174 AACAGAAGAAACACCTAGGAAA-----------------------------ATTTTGAG 1203
              |      ||||||       |||                                 |||||
Db        793 rIleLeuGluThrGluLeuLysAsnValArgAspSerLeuAspGluMetThrGlnLeuAr 813

Qy       1204 GCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGAT 1263
              |        ||||||   :::|||   :::       |||                  |||
Db        813 gAspValLeuGluThrLysAspLysGluAsnGlnThrAlaLeu---LeuGluTyrLysSe 832

Qy       1264 CACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAA 1323
               |||      ::::::|||   |||:::||||||
Db        832 rThrIleHisLysGlnGluAspSerIleLysThr-------------------------- 843

Qy       1324 TAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAAC 1383
                          ||||||||||||          :::::::::
Db        844 ------------LeuGluLysGlyLeuGluThrIleLeuSer---------------Gl 854

Qy       1384 ATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTT 1443
                        ||||||          ::::::          :::   :::|||:::|||
Db        854 nLysLysLysAlaGluAspGlyIleAsnLysMetGly---LysAspLeuPheAlaLeuSe 873

Qy       1444 TGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGC 1503
                                      |||       ::::::::::::   :::   ::
Db        873 rArgGluMetGlnAlaValGluGluAsnCysLysAsnLeuGlnLysGluLysAspLysSe 893

Qy       1504 T---------------------ACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATAC 1542
              :                     ||||||   :::   :::       |||          ||
Db        893 rAsnValAsnHisGlnLysGluThrLysSerLeuLysGluAspIleAlaAlaLysIleTh 913

Qy       1543 GTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGA 1602
              |                    :::||||||   ||||||       :::|||::
Db        913 r--------------------------GluIleLysAlaIleAsnGluAsnLeuGluGl 924

Qy       1603 TCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGG 1662
              :                   :::      :::|||   |||      |||
Db        924 uMetLysIleGlnCysAsnAsnLeuSerLysGluLysGluHisIleSer----------- 940

Qy       1663 GAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAA 1722
```





```
Qy       3469 GGAAGTAGCCACACTGAAACATCAACACCAGGTG------AAGGAAAATAAATACTTTGA 3522
Db       1565 sAspIleGluGluLeuLysSerLysLeuArgIleGluAlaLysSerGlySerGluLeuGl 1585

Qy       3523 GGACATTAAG----------ATTTTACAAGAAAAG---------AATGCTGAACTTCA 3561
Db       1585 uThrValLysGlnGluLeuAsnAsnAlaGlnGluLysIleArgIleAsnAlaGluGluAs 1605

Qy       3562 AATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCA 3621
Db       1605 nThrValLeuLysSerLysLeuGluAspIleGluArgGluLeuLysAspLysGlnAlaGl 1625

Qy       3622 GCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db       1625 uIleLysSerAsnGlnGluGluLysGluLeuLeuThrSerArgLeuLysGlu 1642
```

RESULT 6
CENE_HUMAN
ID CENE_HUMAN STANDARD; PRT; 2663 AA.
AC Q02224;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Centromeric protein E (CENP-E protein).
GN CENPE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93024922; PubMed=1406971;
RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;
RT "CENP-E is a putative kinetochore motor that accumulates just before
RT mitosis.";
RL Nature 359:536-539(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95196755; PubMed=7889940;
RA Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;
RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed
RT microtubule motor.";
RL EMBO J. 14:918-926(1995).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98437347; PubMed=9763420;
RA Chan G.K.T., Schaar B.T., Yen T.J.;
RT "Characterization of the kinetochore binding domain of CENP-E reveals
RT interactions with the kinetochore proteins CENP-F and hBUBR1.";
RL J. Cell Biol. 143:49-63(1998).
RN [4]
RP FARNESYLATION.
RX MEDLINE=20459117; PubMed=10852915;
RA Ashar H.R., James L., Gray K., Carr D., Black S., Armstrong L.,
RA Bishop W.R., Kirschmeier P.;
RT "Farnesyl transferase inhibitors block the farnesylation of CENP-E
RT and CENP-F and alter the association of CENP-E with the
RT microtubules.";
RL J. Biol. Chem. 275:30451-30457(2000).
CC -!- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE
CC KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE
CC OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT
CC AND/OR SPINDLE ELONGATION.
CC -!- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.
CC -!- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING
CC CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS
CC QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.
CC -!- SIMILARITY: Belongs to the kinesin-like protein family.
CC --

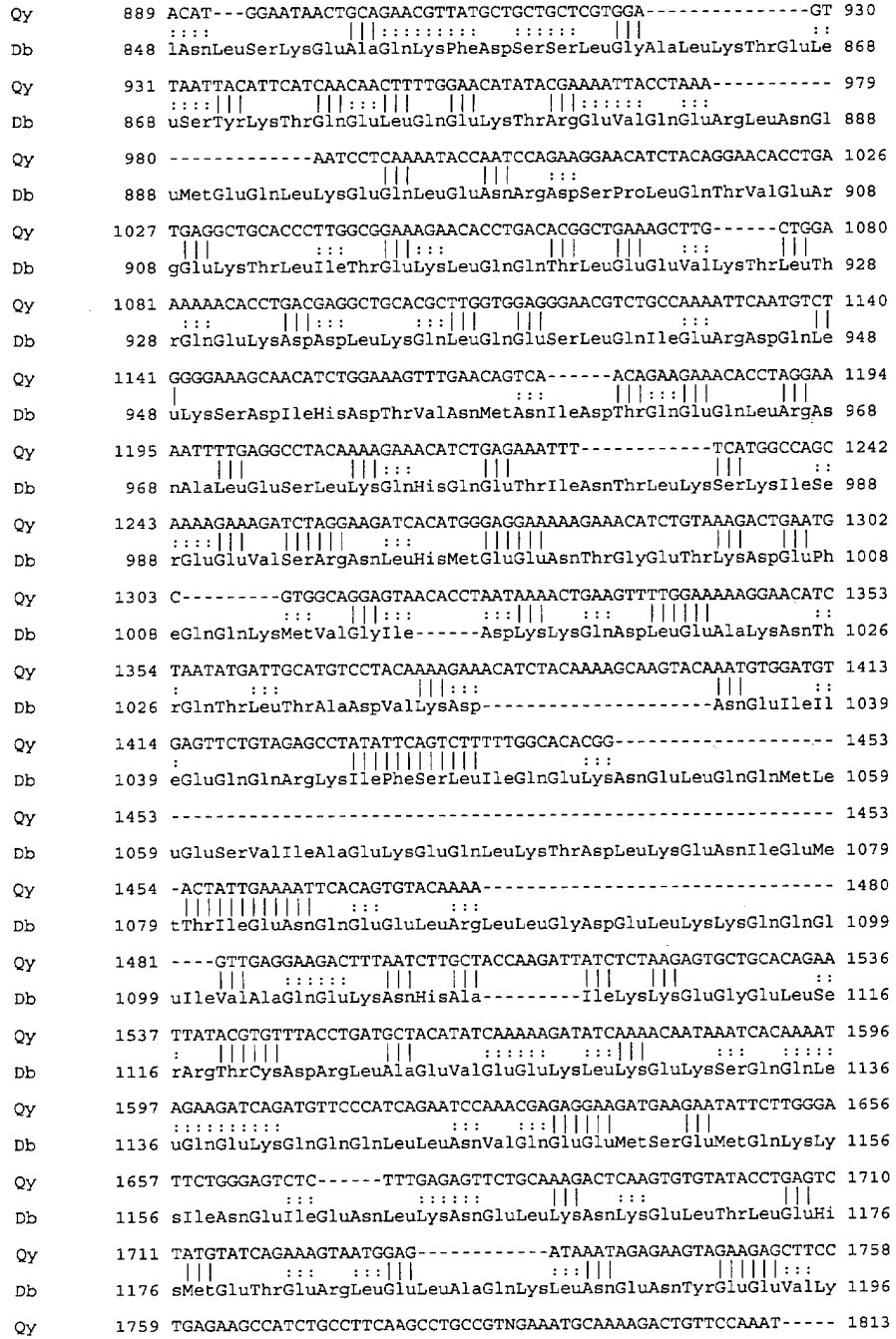
CC --
DR EMBL; Z15005; CAA78727.1; -.
DR PIR; S28261; S28261.
DR HSSP; P17119; 3KAR.
DR Genew; HGNC:1856; CENPE.
DR GK; Q02224; -.
DR MIM; 117143; -.
DR GO; GO:0005699; C:kinetochore; TAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0008350; F:kinetochore motor activity; TAS.
DR GO; GO:0000067; P:DNA replication and chromosome cycle; TAS.
DR GO; GO:0007079; P:mitotic chromosome movement; TAS.
DR GO; GO:0007080; P:mitotic metaphase plate congression; TAS.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISc; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;
KW Cell cycle; Centromere; Lipoprotein; Prenylation.
FT DOMAIN 1 335 KINESIN-MOTOR.
FT DOMAIN 336 2471 COILED COIL (POTENTIAL).
FT DOMAIN 2472 2663 GLOBULAR (POTENTIAL).
FT NP_BIND 86 93 ATP (BY SIMILARITY).
FT LIPID 2660 2660 S-farnesyl cysteine.
SQ SEQUENCE 2663 AA; 312087 MW; CEFC13880C8C8CB8 CRC64;

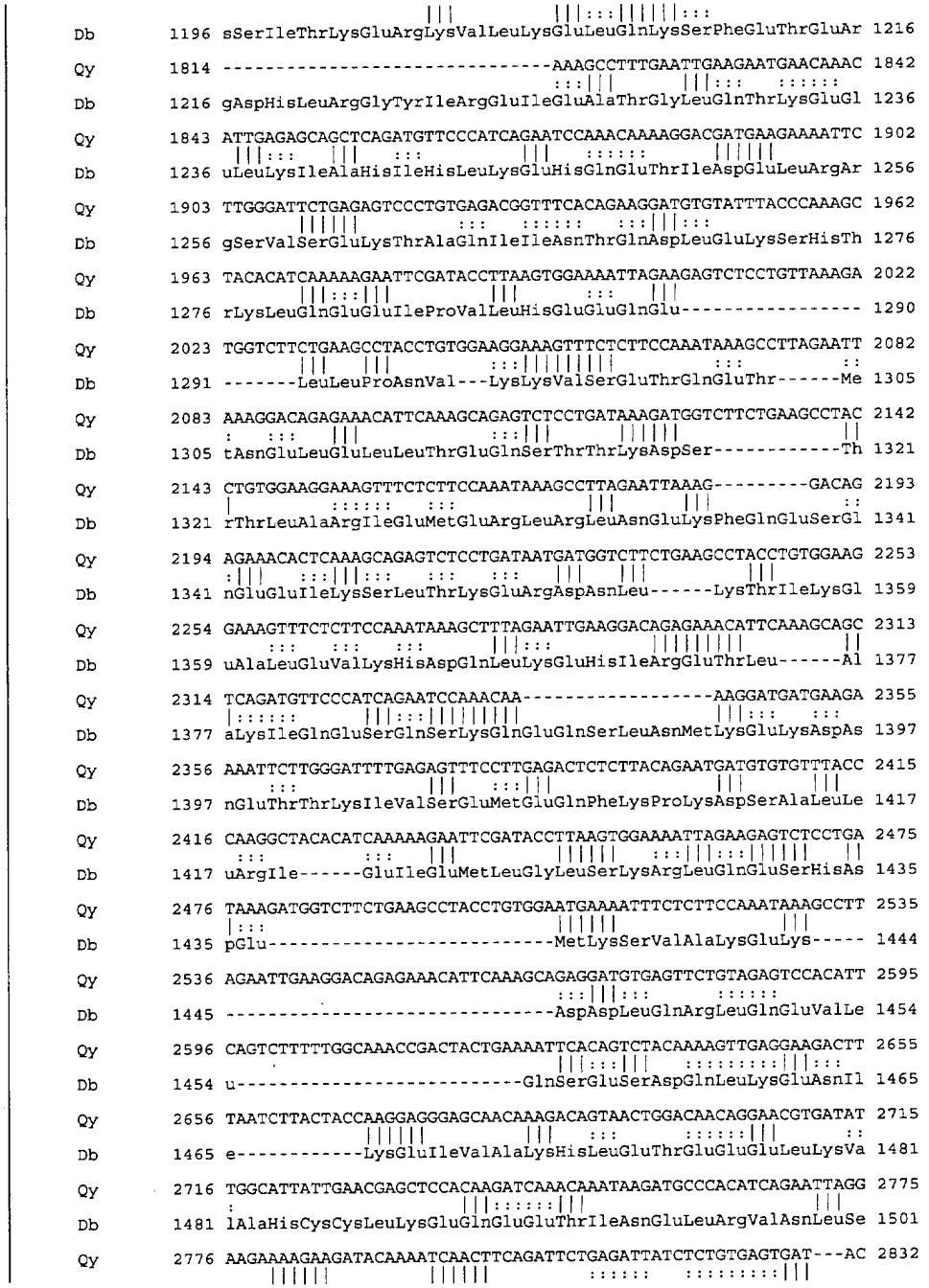
Alignment Scores:

Pred. No.:	2.46e-09	Length:	2663
Score:	330.00	Matches:	250
Percent Similarity:	37.53%	Conservative:	224
Best Local Similarity:	19.79%	Mismatches:	494
Query Match:	5.12%	Indels:	295
DB:	1	Gaps:	50

US-09-602-362E-26 (1-3673) x CENE_HUMAN (1-2663)

```
Qy        478 AATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAA 537
Db        702 SerLeuIleAspGlyLysValProLysAspLeuLeuCysAsnLeuGluLeuGlu-GlyLy 721

Qy        538 GCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGG 597
Db        721 s--------------IleThrAspLeuGlnLysGluLeuAsn---LysGluValGluGl 735

Qy        598 CAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCT 657
Db        735 uAsnGluAlaLeuArg--------------GluGluValIleLeuLeuSerGluLeuLy 750

Qy        658 GTCCTATGGTGCAGTCATC--------------GAGGTGCAAAACAAGGCTAGCCTCAC 702
Db        750 sSerLeuProSerGluValGluArgLeuArgLysGluIleGlnAspLysSerGluGluLe 770

Qy        703 ACCCCTTTTACTGGCC-------------------------ATACAGAAAAGAAGCAA 735
Db        770 uHisIleIleThrSerGluLysAspLysLeuPheSerGluValValHisLysGluSerAr 790

Qy        736 G-------------------------CAAACTGTGGAATTTTTACTAACAAAAAATGC 768
Db        790 gValGlnGlyLeuLeuGluGluIleGlyLysThrLysAspAspLeuAlaThrThrGlnSe 810

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
Db        810 rAsnTyrLysSerThrAspGlnGluPheGlnAsnPheLysThrLeuHisMet------As 828

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
Db        828 pPheGluGlnLysTyrLysMetValLeuGluGluAsnGluArgMetAsnGlnGluIleVa 848
```



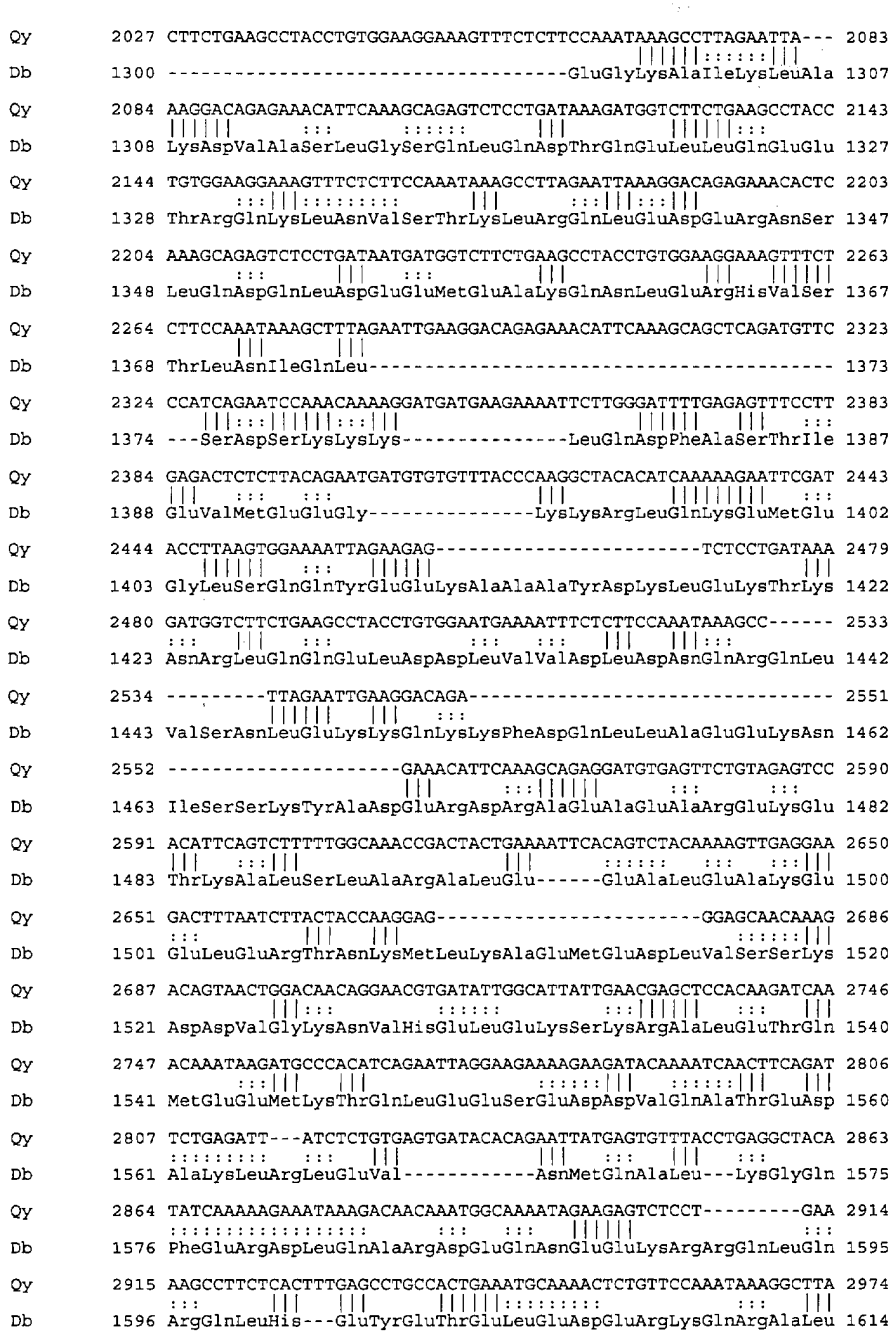


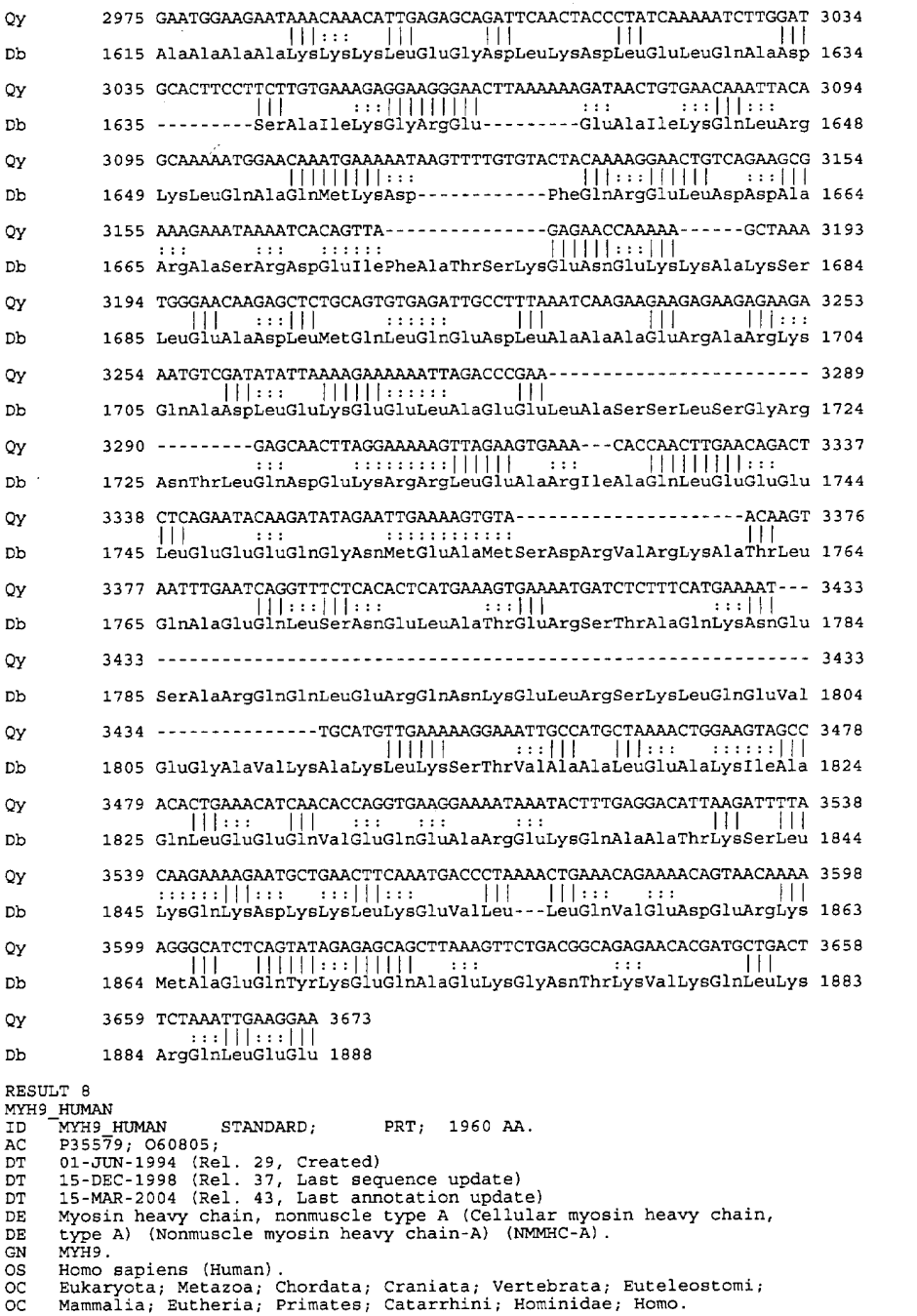
```
Db       1501 rGluLysGluThrGluIleSerThrIleGlnLysGlnLeuGluAlaIleAsnAspLysLe 1521

Qy       2833 ACAGAAT--------------TATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAAT 2877
               |||||              |||||      |||           |||:::||
Db       1521 uGlnAsnLysIleGlnGluIleTyrGlu---LysGluGluGlnLeuAsnIleLysGlnIl 1540

Qy       2878 AAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGC 2937
               |          :::   |||     :::         |||   :::
Db       1540 eSerGluValGlnGluAsnValAsnGluLeuLysGlnPheLysGluHisArgLysAlaLy 1560

Qy       2938 CACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAA-------- 2989
                        |||:::   :::|||   ||||||      |||:::
Db       1560 sAspSerAlaLeuGlnSerIleGluSerLysMetLeuGluLeuThrAsnArgLeuGlnGl 1580

Qy       2990 ----------------CAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGA 3033
                             |||   :::      :::         :::   ::::::   ::
Db       1580 uSerGlnGluGluIleGlnIleMetIleLysGluLysGluGluMetLysArgValGlnGl 1600

Qy       3034 TGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTAC 3093
               :||||||         :::       |||:::       |||:::|||   ::::::|||
Db       1600 uAlaLeu---------GlnIleGluArgAspGlnLeuLysGluAsnThrLysGluIleVa 1617

Qy       3094 AGCAAAAATGGAACAAATGAAAAATAAG---------TTTTGTGTACTA----------- 3133
                |||||||||::::::   :::   |||         |||       :::
Db       1617 lAlaLysMetLysGluSerGlnGluLysGluTyrGlnPheLeuLysMetThrAlaValAs 1637

Qy       3134 -------CAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAA 3186
                    |||:::::::::   |||   :::   :::|||   |||   |||   |||||
Db       1637 nGluThrGlnGluLysMetCysGluIleGluHisLeuLysGluGlnPheGluThrGlnLy 1657

Qy       3187 AGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAA 3246
               |         |||                                 |||   |||   |||
Db       1657 sLeuAsnLeuGlu-----------------------------AsnIleGluThrGluAs 1667

Qy       3247 GAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAA 3306
                  |||         ||||||   |||   :::                 :::   |||::
Db       1667 nIleArgLeuThrGlnIleLeuHisGluAsnLeuGluGluMetArgSerValThrLysGl 1687

Qy       3307 GTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAA-- 3364
                :   :::           :::|||:::||||||:::::::::      :::||||||
Db       1687 uArgAspAspLeuArgSerValGluGluThrLeuLysValGluArgAspGlnLeuLysGl 1707

Qy       3365 -------------AGTGTAACAAGTAATTTGAAT--------------CAGGTTTCTCA 3396
                          ::::::|||   :::|||                 ::::::   ||
Db       1707 uAsnLeuArgGluThrIleThrArgAspLeuGluLysGlnGluGluLeuLysIleValHi 1727

Qy       3397 CACTCATGAAAGTGAAAAT---------------------------------------- 3415
               |   |||       |||:::
Db       1727 sMetHisLeuLysGluHisGlnGluThrIleAspLysLeuArgGlyIleValSerGluLy 1747

Qy       3416 ----------------------------GATCTCTTTCATGAAAATTGCATGTTGAAAAA 3447
                                         ||||||   |||   |||       ||||||
Db       1747 sThrAsnGluIleSerAsnMetGlnLysAspLeuGluHisSerAsnAspAlaLeuLysAl 1767

Qy       3448 GGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGA 3507
                :::       ::::::   |||:::      :::   |||   |||       :::::
Db       1767 aGlnAspLeuLysIleGlnGluGluLeu---ArgIleAlaHisMetHisLeuLysGluGl 1786

Qy       3508 AAATAAATACTTTGAGGACATT------------------------------------- 3529
               :   :::      :::   :::
Db       1786 nGlnGluThrIleAspLysLeuArgGlyIleValSerGluLysThrAspLysLeuSerAs 1806

Qy       3530 -------AAGATTTTACAAGAAAAGAATGCTGAACTTCAA-------------------- 3562
                     |||   |||:::      ||||||:::||||||
Db       1806 nMetGlnLysAspLeuGluAsnSerAsnAlaLysLeuGlnGluLysIleGlnGluLeuLy 1826

Qy       3563 -------------------ATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGC 3603
                              :::|||||||||                        |||:::
Db       1826 sAlaAsnGluHisGlnLeuIleThrLeuLysLysAspValAsnGluThrGlnLysLysVa 1846

Qy       3604 ATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAA 3663
                |||:::      ||||||||||||          :::   ::::::      |||||
Db       1846 lSerGlu---MetGluGlnLeuLysLysGlnIleLysAspGlnSerLeuThrLeuSerLy 1865

Qy       3664 ATTGAAG 3670
                 ||||:::
Db       1865 sLeuGlu 1867
```

RESULT 7
MYHB_MOUSE
ID MYHB_MOUSE STANDARD; PRT; 1972 AA.
AC O08638; O08639; Q62462; Q64195;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, smooth muscle isoform (SMMHC).
GN MYH11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Uterus;
RX MEDLINE=97242182; PubMed=9125171;
RA Hasegawa K., Arakawa E., Oda S., Matsuda Y.;
RT "Molecular cloning and expression of murine smooth muscle myosin heavy
RT chains.";
RL Biochem. Biophys. Res. Commun. 232:313-316(1997).
RN [2]
RP SEQUENCE OF 1-368 FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=95008063; PubMed=7923625;
RA Miano J.M., Cserjesi P., Ligon K.L., Periasamy M., Olson E.N.;
RT "Smooth muscle myosin heavy chain exclusively marks the smooth muscle
RT lineage during mouse embryogenesis.";
RL Circ. Res. 75:803-812(1994).
RN [3]
RP SEQUENCE OF 1-126 FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=96172919; PubMed=8593698;
RA Suzuki T., Kim H.S., Kurabayashi M., Hamada H., Fujii H., Aikawa M.,
RA Watanabe M., Watanabe N., Sakomura Y., Yazaki Y., Nagai R.;
RT "Preferential differentiation of P19 mouse embryonal carcinoma cells
RT into smooth muscle cells. Use of retinoic acid and antisense against
RT the central nervous system-specific POU transcription factor Brn-2.";
RL Circ. Res. 78:395-404(1996).
CC -!- FUNCTION: Muscle contraction.
CC -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC and 2 regulatory light chain subunits (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O08638-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O08638-2; Sequence=VSP_003346;
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be
CC split further into 2 globular subfragments (S1) and 1 rod-shaped
CC subfragment (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC ---

CC ---

```
DR   EMBL; D85923; BAA19690.1; -.
DR   EMBL; D85924; BAA19691.1; -.
DR   EMBL; L25860; AAA67552.1; -.
DR   EMBL; S81516; AAB36168.1; -.
DR   PIR; I52863; I52863.
DR   PIR; JC5420; JC5420.
DR   PIR; JC5421; JC5421.
DR   HSSP; P10587; 1BR2.
DR   MGD; MGI:102643; Myh11.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   Calmodulin-binding; ATP-binding; Methylation; Multigene family;
KW   Alternative splicing.
FT   DOMAIN       1    785       MYOSIN HEAD-LIKE.
FT   DOMAIN     786    815       IQ.
FT   DOMAIN     844   1934       COILED COIL (POTENTIAL).
FT   DOMAIN    1935   1972       CARBOXYL-TERMINAL.
FT   NP_BIND    178    185       ATP (POTENTIAL).
FT   DOMAIN     661    683       ACTIN-BINDING (BY SIMILARITY).
FT   DOMAIN     762    776       ACTIN-BINDING (BY SIMILARITY).
FT   MOD_RES    129    129       METHYLATION (TRI-) (POTENTIAL).
FT   VARSPLIC  1930   1972       RGNEASFVPSRRAGGRRVIENTDGSEEEMDARDSDFNGTKA
FT                               SE -> GPPPQETSQ (in isoform 2).
FT                               /FTId=VSP_003346.
FT   CONFLICT   126    126       N -> D (IN REF. 3).
FT   CONFLICT   161    161       A -> V (IN REF. 2).
FT   CONFLICT   189    189       Q -> K (IN REF. 2).
SQ   SEQUENCE  1972 AA;  227026 MW;  A1398E3F5B11F15A CRC64;

Alignment Scores:
Pred. No.:              4.53e-09      Length:         1972
Score:                  324.50        Matches:        227
Percent Similarity:     37.07%        Conservative:   190
Best Local Similarity:  20.18%        Mismatches:     421
Query Match:            5.03%         Indels:         288
DB:                     1             Gaps:           40

US-09-602-362E-26 (1-3673) x MYHB_MOUSE (1-1972)

Qy        951 TTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACAT 1010
              :::|||   ::::::                :::|||::::::     |||:::|||
Db        834 TrpTrpArgLeuPheThrLysValLysProLeuLeuGlnValThrArgGlnGluGluGlu 853

Qy       1011 CTACAG---------GAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACA 1061
              :::|||         |||   :::   ::::::        :::::::::      |||
Db        854 MetGlnAlaLysGluGluGluMetGlnLysIleThrGluArgGlnGlnLysAlaGluThr 873

Qy       1062 CGGCTGAAAGCTTGCTGGAAAAAACACCTG---------------------------- 1091
                 ||||||         :::||||||
Db        874 GluLeuLysGluLeuGluGlnLysHisThrGlnLeuAlaGluGluLysThrLeuLeuGln 893

Qy       1092 ------------------ACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTC 1133
                               |||   |||:::|||        |||        :::
Db        894 GluGlnLeuGlnAlaGluThrGluLeuTyrAlaGluSerGluGluMetArgValArgLeu 913

Qy       1134 AATGTCTGGGGAAAGCAACATCTGGAA------------------------------ 1160
                        ||||||   ||||||
Db        914 AlaAla---LysLysGlnGluLeuGluGluIleLeuHisGluMetGluAlaArgLeuGlu 932

Qy       1161 --------------------AGTTTGAACAGTCAACAGAAGAAACA-CCTAGGAAAATT 1198
                                   |||   :::::::::||||||      :::::::::

Db        933 GluGluGluAspArgArgGlnGlnLeuGlnAlaGluArgLysLysMetAlaGlnGlnMet 952

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
              |||         ::::::      |||:::   :::        |||   :::
Db        953 LeuAspLeuGluGluGlnLeuGluGluGluGluAlaAlaArgGlnLysLeuGlnLeuGlu 972

Qy       1259 AAGATCACATGGGAG---------------------------------------------- 1273
              |||:::|||   |||
Db        973 LysValThrAlaGluAlaLysIleLysLysLeuGluAspAspIleLeuValMetAspAsp 992

Qy       1274 ---------------GAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
                            ||||||      :::      |||   |||:::   :::|||
Db        993 GlnAsnSerLysLeuSerLysGluArgLysLeuLeuGluGluArgValSerAspLeuThr 1012

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
                 |||       |||       ||||||          |||:::           :::|||
Db       1013 ThrAsnLeuAlaGluGluGluGluLys---AlaLysAsnLeuThrLysLeuLysSerLys 1031

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTG---------------------------- 1408
                    |||       :::       |||
Db       1032 HisGluSerMetIleSerGluLeuGluValArgLeuLysLysGluGluLysSerArgGln 1051

Qy       1409 --------------------------------GATGTGAGTTCTGTA---GAGCCTATA 1432
                                               |||   |||         |||   |||
Db       1052 GluLeuGluLysLeuLysArgLysLeuGluGlyAspAlaSerAspPheHisGluGlnIle 1071

Qy       1433 TTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGAC 1492
                 |||                                 |||       |||   ||||||:::
Db       1072 AlaAspLeuGlnAlaGlnIleAlaGluLeuLysMetGlnLeuAlaLysLysGluGluGlu 1091

Qy       1493 TTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCT 1552
                    |||         :::   :::       ||||||                |||
Db       1092 LeuGlnAlaAlaLeuAlaArgLeuAspGluGluIleAlaGlnLysAsnAsnAlaLeu--- 1110

Qy       1553 GATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTC 1612
                       |||   |||:::   :::       |||   |||
Db       1111 --------------LysLysIleArgGluLeuGluGlyHisIleSerAspLeuGlnGlu 1125

Qy       1613 CCATCAGAATCCAAA---------CGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGG 1663
                 :::|||:::          |||   :::   |||:::          |||   |||
Db       1126 AspLeuAspSerGluArgAlaAlaArgAsnLysAlaGluLysGlnLysArgAspLeuGly 1145

Qy       1664 AGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTG------TGTATACCTGAGTCTATGTAT 1717
                    |||   :::   |||||:::::::         :::       :::
Db       1146 GluGluLeuGlu---AlaLeuLysThrGluLeuGluAspThrLeuAspSerThrAlaThr 1164

Qy       1718 CAGAAAGTAATGGAGATAAATAGAGAA------------------------------- 1744
                 |||:::   :::          ||||||
Db       1165 GlnGlnGluLeuArgAlaLysArgGluGlnGluValThrValLeuLysLysAlaLeuAsp 1184

Qy       1744 ------------------------------------------------------------ 1744

Db       1185 GluGluThrArgSerHisGluAlaGlnValGlnGluMetArgGlnLysHisThrGlnAla 1204

Qy       1745 GTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTN------------GAA 1792
              |||||||||||||   |||:::         ||||||   |||                :::
Db       1205 ValGluGluLeuThrGluGlnLeuGluGlnPheLysArgAlaLysLysAlaAsnLeuAspLys 1224

Qy       1793 ATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCA 1852
                 ::::::|||:::       :::       :::|||       |||       :::
Db       1225 SerLysGlnThrLeuGluLysGluAsnAlaAspLeuAlaGlyGluLeuArgValLeuGly 1244

Qy       1853 GCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCT 1912
                            |||   |||:::|||       |||             |||
Db       1245 GlnAlaLysGlnGluValGluHisLysLysLysLysLeuGluValGlnLeuGlnAspLeu 1264

Qy       1913 GAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACAT--- 1969
              :::|||   |||           :::           |||             |||
Db       1265 GlnSerLysCysSerAspGlyGluArgAlaArgAlaGluLeuSerAspLysValHisLys 1284

Qy       1970 ---CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGT 2026
                  |||   |||   ::::::::::|||   |||   |||:::
Db       1285 LeuGlnAsnGluValGluSerValThrGlyMetLeuAsnGluAla--------------- 1299
```





RESULT 8
MYH9_HUMAN
ID MYH9_HUMAN STANDARD; PRT; 1960 AA.
AC P35579; O60805;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, nonmuscle type A (Cellular myosin heavy chain,
DE type A) (Nonmuscle myosin heavy chain-A) (NMMHC-A).
GN MYH9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20057165; PubMed=10591208;
RA   Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
RA   Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA   Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA   Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA   Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA   Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA   Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA   Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA   Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA   Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA   Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA   Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA   Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA   Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA   Mcclay J., Mclaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA   Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA   Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA   Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA   Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA   Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA   Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA   Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA   Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA   Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA   Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA   Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA   Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA   Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA   Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA   Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RA   Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA   Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA   Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA   Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA   Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA   Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA   Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA   Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA   Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA   O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA   Khan A.S., Lane L., Tilahun Y., Wright H.;
RT   "The DNA sequence of human chromosome 22.";
RL   Nature 402:489-495(1999).
RN   [2]
RP   SEQUENCE OF 1-1337 FROM N.A.
RX   MEDLINE=92003925; PubMed=1912569;
RA   Toothaker L.E., Gonzalez D.A., Tung N., Lemons R.S., le Beau M.M.,
RA   Arnaout M.A., Clayton L.K., Tenen D.G.;
RT   "Cellular myosin heavy chain in human leukocytes: isolation of 5'
RT   cDNA clones, characterization of the protein, chromosomal
RT   localization, and upregulation during myeloid differentiation.";
RL   Blood 78:1826-1833(1991).
RN   [3]
RP   SEQUENCE OF 1-715 FROM N.A.
RX   MEDLINE=91316803; PubMed=1860190;
RA   Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
RA   Gdula D., Adelstein R.S., Weir L.;
RT   "Human nonmuscle myosin heavy chains are encoded by two genes located
RT   on different chromosomes.";
RL   Circ. Res. 69:530-539(1991).
RN   [4]
RP   SEQUENCE OF 714-1960 FROM N.A.
RX   MEDLINE=90138958; PubMed=1967836;
RA   Saez C.G., Myers J.C., Shows T.B., Leinwand L.A.;
RT   "Human nonmuscle myosin heavy chain mRNA: generation of diversity
RT   through alternative polyadenylylation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:1164-1168(1990).
RN   [5]
RP   VARIANT DFNA17 HIS-705.
RX   MEDLINE=20489856; PubMed=11023810;
RA   Lalwani A.K., Goldstein J.A., Kelley M.J., Luxford W., Castelein C.M.,
RA   Mhatre A.N.;
RT   "Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in
RT   nonmuscle myosin MYH9.";
RL   Am. J. Hum. Genet. 67:1121-1128(2000).
RN   [6]
RP   VARIANTS MHA/FTNS/SBS LYS-93; CYS-702; CYS-1165; HIS-1424 AND
RP   LYS-1841.
RX   MEDLINE=20428192; PubMed=10973259;
RA   Seri M., Cusano M., Gangarossa S., Caridi G., Bordo D., Lo Nigro C.,
RA   Ghiggeri G.M., Ravazzolo R., Savino M., Del Vecchio M., d'Apolito M.,
RA   Iolascon A., Zelante L.L., Savoia A., Balduini C.L., Noris P.,
RA   Magrini U., Belletti S., Heath K.E., Babcock M., Glucksman M.J.,
RA   Aliprandis E., Bizzaro N., Desnick R.J., Martignetti J.A.;
RT   "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and
RT   Sebastian syndromes.";
RL   Nat. Genet. 26:103-105(2000).
RN   [7]
RP   VARIANTS MHA ILE-1155 AND LYS-1841.
RX   MEDLINE=20428193; PubMed=10973260;
RA   Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.;
RT   "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in
RT   May-Hegglin anomaly.";
RL   Nat. Genet. 26:106-108(2000).
CC   -!- FUNCTION: Cellular myosin appears to play a role in cytokinesis,
CC       cell shape, and specialized functions such as secretion and
CC       capping.
CC   -!- SUBUNIT: Myosin is a hexameric protein that consists of 2 heavy
CC       chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2
CC       regulatory light chain subunits (MLC-2).
CC   -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC       cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC       characteristic for alpha-helical coiled coils.
CC   -!- DISEASE: Defects in MYH9 are the cause of May-Hegglin anomaly
CC       (MHA) [MIM:155100]. MHA is an autosomal dominant
CC       macrothrombocytopenia characterized by thrombocytopenia, giant
CC       platelets and leukocyte inclusions.
CC   -!- DISEASE: Defects in MYH9 are the cause of Fechtner syndrome (FTNS)
CC       [MIM:153640]. FTNS is an autosomal dominant macrothrombocytopenia
CC       characterized by thrombocytopenia, giant platelets and leukocyte
CC       inclusions. With additional alport-like clinical features of
CC       sensorineural deafness, cataracts and nephritis.
CC   -!- DISEASE: Defects in MYH9 are the cause of Sebastian syndrome (SBS)
CC       [MIM:605249]. SBS is an autosomal dominant macrothrombocytopenia
CC       characterized by thrombocytopenia, giant platelets and leukocyte
CC       inclusions.
CC   -!- DISEASE: Defects in MYH9 are the cause of autosomal dominant
CC       nonsyndromic sensorineural deafness 17 (DFNA17) [MIM:603622].
CC       DFNA17 is characterized by progressive hearing impairment and
CC       cochleosaccular degeneration.
CC   -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC   -!- SIMILARITY: Contains 1 IQ domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z82215; -; NOT_ANNOTATED_CDS.
DR   EMBL; M81105; AAA59888.1; -.
DR   EMBL; M69180; AAA61765.1; -.
DR   EMBL; M31013; AAA36349.1; -.
DR   PIR; A61231; A61231.
DR   HSSP; P10587; 1BR2.
DR   Genew; HGNC:7579; MYH9.
DR   MIM; 160775; -.
DR   MIM; 155100; -.
DR   MIM; 153640; -.
DR   MIM; 605249; -.
```

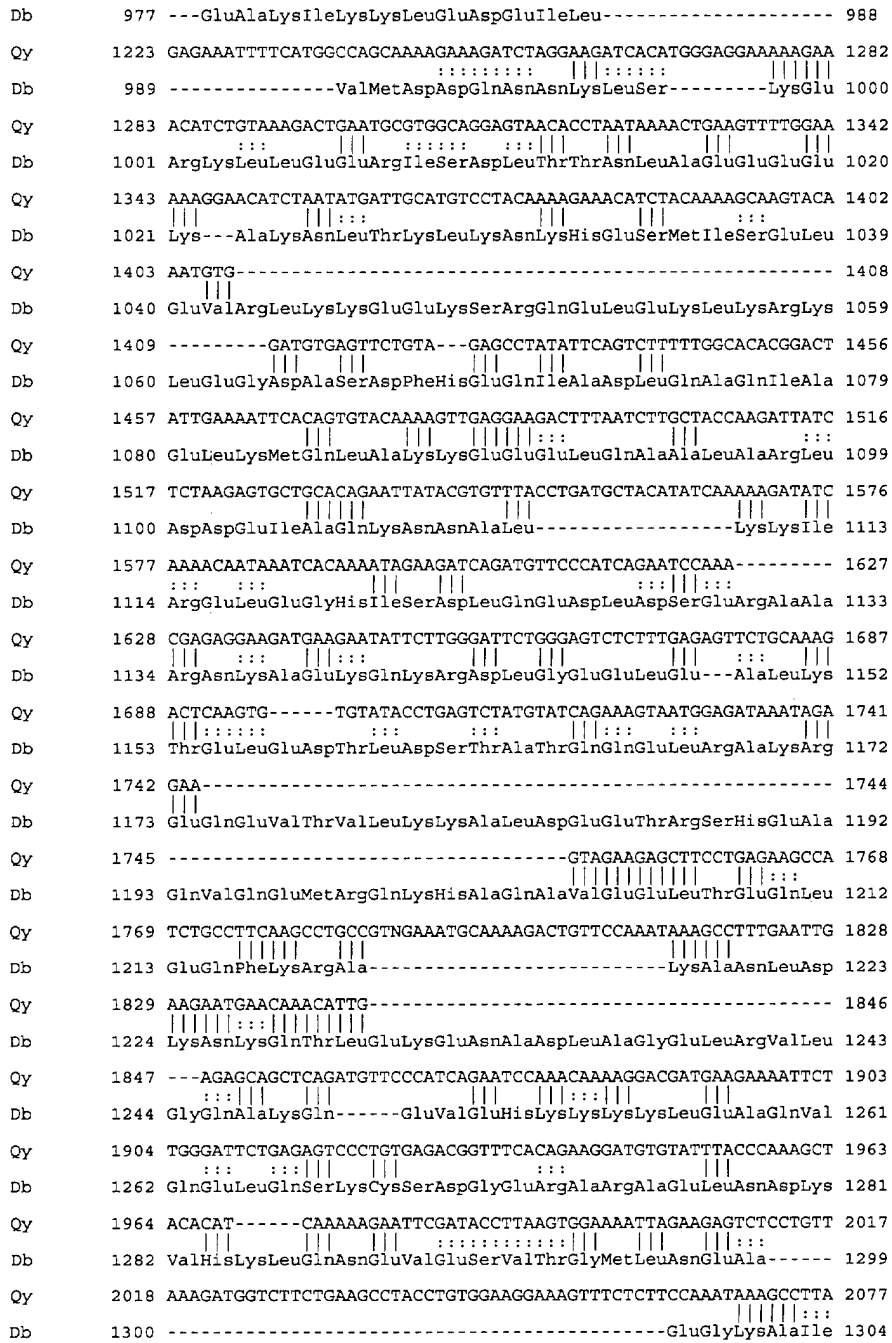
```
DR   MIM; 603622; -.
DR   GO; GO:0005860; C:non-muscle myosin; TAS.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR002017; Spectrin.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; ATP-binding; Calmodulin-binding; Actin-binding; Coiled coil;
KW   Multigene family; Disease mutation; Deafness.
FT   DOMAIN        1    778       MYOSIN HEAD-LIKE.
FT   DOMAIN      779    808       IQ.
FT   DOMAIN      837   1926       COILED COIL (POTENTIAL).
FT   NP_BIND     174    181       ATP (POTENTIAL).
FT   DOMAIN      654    676       ACTIN-BINDING.
FT   VARIANT      93     93       N -> K (IN MHA).
FT                                /FTId=VAR_010791.
FT   VARIANT     702    702       R -> C (IN FTNS).
FT                                /FTId=VAR_010792.
FT   VARIANT     705    705       R -> H (IN DFNA17).
FT                                /FTId=VAR_010793.
FT   VARIANT    1155   1155       T -> I (IN MHA).
FT                                /FTId=VAR_010794.
FT   VARIANT    1165   1165       R -> C (IN SBS).
FT                                /FTId=VAR_010795.
FT   VARIANT    1424   1424       D -> H (IN FTNS).
FT                                /FTId=VAR_010796.
FT   VARIANT    1841   1841       E -> K (IN MHA).
FT                                /FTId=VAR_010797.
FT   CONFLICT     53     55       EAI -> RGH (IN REF. 3).
FT   CONFLICT    660    660       T -> S (IN REF. 3).
FT   CONFLICT    869    869       T -> M (IN REF. 4).
FT   CONFLICT    931    931       C -> Y (IN REF. 4).
FT   CONFLICT   1240   1241       KG -> GR (IN REF. 4).
FT   CONFLICT   1350   1350       E -> EE (IN REF. 2).

Alignment Scores:
Pred. No.:                 8.44e-09        Length:          1960
Score:                     319.00          Matches:         219
Percent Similarity:        38.23%          Conservative:    187
Best Local Similarity:     20.62%          Mismatches:      434
Query Match:               4.95%           Indels:          222
DB:                        1               Gaps:            40

US-09-602-362E-26 (1-3673) x MYH9_HUMAN (1-1960)

Qy        854 CTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTAT 913
Db        920 LeuGluAlaArgValGluGluGluGluGluArgCysGlnHisLeuGlnAlaGluLysLys 939

Qy        914 GCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTA 973
Db        940 LysMetGlnGlnAsnIleGlnGluLeuGluGluGlnLeuGluGluGlu------------ 955

Qy        974 CCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCT 1033
Db        956 ----------------------------------------------------GluSerAla 958

Qy       1034 GCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGAC 1093
Db        959 ArgGlnLysLeuGlnLeuGluLysValThrThrGluAlaLysLeuLysLysLeuGluGlu 978

Qy       1094 GAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACA 1153
Db        979 GluGlnIle---IleLeuGluAspGlnAsnCysLysLeuAlaLysGluLysLysLeuLeu 997

Qy       1154 TCTGGAAAGTTTGAACAGTCAACA--------------GAAGAAACACCTAGGAAAATT 1198
Db        998 GluAspArgIleAlaGluPheThrThrAsnLeuThrGluGluGluGluLysSerLysSer 1017

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
Db       1018 LeuAlaLysLeuLysAsnLysHisGluAlaMetIleThrAspLeuGluGluArgLeuArg 1037

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
Db       1038 Arg---------GluGluLysGlnArgGlnGluLeuGluLys--------------Thr 1049

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCA------------ 1366
Db       1050 ArgArgLysLeuGluGlyAspSerThrAspLeuSerAspGlnIleAlaGluLeuGlnAla 1069

Qy       1367 --------------------------TGTCCTACAAAAGAAACATCTACAAAAGCAAGT 1399
Db       1070 GlnIleAlaGluLeuLysMetGlnLeuAlaLysLysGluGluGluLeuGlnAlaAlaLeu 1089

Qy       1400 ACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATT 1459
Db       1090 AlaArgValGluGluGluAlaAlaGlnLysAsnMetAlaLeuLysLysIleArgGluLeu 1109

Qy       1460 GAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCT 1519
Db       1110 Glu---SerGlnIleSerGluLeuGlnGluAspLeuGlu------------------Ser 1122

Qy       1520 AAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAA 1579
Db       1123 GluArgAlaSerArgAsnLysAlaGluLysGlnLysArgAspLeuGlyGluGluLeuGlu 1142

Qy       1580 ACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCA------------------GAA 1621
Db       1143 AlaLeuLysThrGluLeuGluAspThrLeuAspSerThrAlaAlaGlnGlnGluLeuArg 1162

Qy       1622 TCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCT 1681
Db       1163 SerLysArgGluGlnGluValAsnIle---------LeuLysLysThrLeuGluGluGlu 1179

Qy       1682 GCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGA 1741
Db       1180 AlaLysThrHisGluAlaGlnIleGlnGluMetArgGlnLysHisSerGlnAlaValGlu 1199

Qy       1742 GAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAG 1801
Db       1200 GluLeuAlaGluGlnLeuGluGlnThrLysArgValLysAlaAsnLeuGluLysAlaLys 1219

Qy       1802 ---ACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAG 1858
Db       1220 GlnThrLeuGluAsnGluArgGlyGluLeuAlaAsnGluValLysValLeuGluGlnGly 1239

Qy       1859 ATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGT 1918
Db       1240 LysGlyAspSerGluHisLysArgLysLysValGluAlaGlnLeuGlnGluLeuGlnVal 1259

Qy       1919 CCCTGT---GAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACAT---CAA 1972
Db       1260 LysPheAsnGluGlyGluArgValArgThrGluLeuAlaAspLysValThrLysLeuGln 1279

Qy       1973 AAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTG 2032
Db       1280 ValGluLeuAspAsnValThrGlyLeuLeuSerGlnSerAspSerLysSerSerLysLeu 1299

Qy       2033 AAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGA 2092
Db       1300 ------------------------------------------------ThrLysAspPhe 1303

Qy       2093 GAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGG 2152
Db       1304 SerAlaLeuGluSerGlnLeuGlnAspThrGlnGluLeuLeuGlnGluGluAsnArgGln 1323
```

```
Qy       2153 AAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGAC--------------AGAGAA 2197
Db       1324 LysLeuSerLeuSerThrLysLeuLysGlnValGluAspGluLysAsnSerPheArgGlu 1343

Qy       2198 ACACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAG------CCTACCTGTGGA 2251
Db       1344 GlnLeuGluGluGluGluGluAlaLysHisAsnLeuGluLysGlnIleAlaThrLeuHis 1363

Qy       2252 AGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAA--------------TTGAAGGACAGA 2296
Db       1364 AlaGlnValAlaAspMetLysLysLysMetGluAspSerValGlyCysLeuGluThrAla 1383

Qy       2297 GAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAA------TCCAAACAAAAGGATGAT 2350
Db       1384 GluGluValLysArgLysLeuGlnLysAspLeuGluGlyLeuSerGlnArgHisGluGlu 1403

Qy       2351 GAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAAT------GAT 2404
Db       1404 LysValAlaAlaTyrAspLysLeuGluLysThrLysThrArgLeuGlnGlnGluLeuAsp 1423

Qy       2405 GTGTGTTTACCCAAGGCTACACAT--------------------------------CAA 2431
Db       1424 AspLeuLeuValAspLeuAspHisGlnArgGlnSerAlaCysAsnLeuGluLysLysGln 1443

Qy       2432 AAAGAATTCGAT--------------------ACCTTAAGTGGAAAATTAGAAGAGTCT 2470
Db       1444 LysLysPheAspGlnLeuLeuAlaGluGluLysThrIleSerAlaLysTyrAlaGluGlu 1463

Qy       2471 CCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAA 2530
Db       1464 ArgAspArg----------------------------------------------------Ala 1467

Qy       2531 GCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCC 2590
Db       1468 GluAlaGluAlaArgGluLysGluThrLysAlaLeuSerLeuAlaArgAlaLeuGluGlu 1487

Qy       2591 ACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAA 2650
Db       1488 AlaMetGluGlnLysAlaGluLeuGluArgLeuAsnLysGlnPheArgThrGluMetGlu 1507

Qy       2651 GACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGT 2710
Db       1508 AspLeu---MetSerSerLysAspAspValGlyLysSerVal---------------His 1521

Qy       2711 GATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCAGAA 2770
Db       1522 GluLeuGluLysSerLysArgAlaLeuGluGlnGlnValGluGluMetLysThrGlnLeu 1541

Qy       2771 TTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATT---ATCTCTGTGAGT 2827
Db       1542 GluGluLeuGluAspGluLeuGlnAlaThrGluAspAlaLysLeuArgLeuGluVal--- 1560

Qy       2828 GATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACA 2887
Db       1561 ---------AsnLeuGlnAlaMet---LysAlaGlnPheGluArgAspLeuGln------ 1574

Qy       2888 AATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTT---------------GAG 2932
Db       1575 ---GlyArgAspGluGlnSerGluGluLysLysGlnLeuValArgGlnValArgGlu 1593

Qy       2933 CCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAA 2992
Db       1594 MetGluAlaGluLeuGluAspGluArgLysGlnArgSerMetAlaValAlaAlaArgLys 1613

Qy       2993 ACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCA--------------- 3037
Db       1614 LysLeuGluMetAspLeuLysAspLeuGluAlaHisIleAspSerAlaAsnLysAsnArg 1633

Qy       3038 ------------------------------------CTTCCTTCTTGTGAAAGAGGA 3058
Db       1634 AspGluAlaIleLysGlnLeuArgLysLeuGlnAlaGlnMetLysAspCysMetArgGlu 1653

Qy       3059 AGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCA--------------------- 3097
```

```
Db       1654 LeuAspAspThrArgAlaSerArgGluGluIleLeuAlaGlnAlaLysGluAsnGluLys 1673

Qy       3098 AAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAA 3157
Db       1674 LysLeuLysSerMetGluAlaGluMetIleGlnLeuGlnGluGluLeuAlaAlaAlaGlu 1693

Qy       3158 GAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTG 3217
Db       1694 ArgAlaLysArgGlnAlaGlnGlnGluArgAspGluLeuAlaAspGluIleAlaAsnSer 1713

Qy       3218 ------------AGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATA 3265
Db       1714 SerGlyLysGlyAlaLeuAlaLeu---------GluGluLysArgArg------------ 1726

Qy       3266 TTAAAAGAAAAAATT---AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACAC 3322
Db       1727 LeuGluAlaArgIleAlaGlnLeuGluGluGluLeuGluGluGluGlnGlyAsnThrGlu 1746

Qy       3323 CAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTG 3382
Db       1747 LeuIleAsnAspArgLeuLysLysAlaAsnLeuGlnIleAspGlnIleAsnThrAspLeu 1766

Qy       3383 AAT---CAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATG 3439
Db       1767 AsnLeuGluArgSerHisAlaGlnLysAsnGluAsn------------AlaArgGlnGln 1782

Qy       3440 TTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAG 3499
Db       1783 LeuGluArgGlnAsnLysGluLeuLysValLys---------LeuGlnGluMetGluGly 1799

Qy       3500 GTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTT 3559
Db       1800 ThrValLysSerLysTyrLysAlaSerIleThrAlaLeuGluAlaLysIleAlaGlnLeu 1819

Qy       3560 CAAATGACCCTA---------------------------------------------- 3571
Db       1820 GluGluGlnLeuAspAsnGluThrLysGluArgGlnAlaAlaCysLysGlnValArgArg 1839

Qy       3572 ---------AAACTGAAACAGAAAACAGTA--------------ACAAAAAGGGCATCT 3607
Db       1840 ThrGluLysLysLeuLysAspValLeuLeuGlnValAspAspGluArgArgAsnAlaGlu 1859

Qy       3608 CAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTG 3667
Db       1860 GlnTyrLysAspGlnAlaAspLysAlaSerThrArgLeuLysGlnLeuLysArgGlnLeu 1879

Qy       3668 AAGGAA 3673
Db       1880 GluGlu 1881
```

```
RESULT 9
MYHB_HUMAN
ID   MYHB_HUMAN     STANDARD;      PRT;  1972 AA.
AC   P35749; O00396; O94944; P78422;
DT   01-JUN-1994 (Rel. 29, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Myosin heavy chain, smooth muscle isoform (SMMHC).
GN   MYH11 OR KIAA0866.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99425270; PubMed=10493829;
RA   Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA   Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA   Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA   Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT   "Genome duplications and other features in 12 Mb of DNA sequence from
RT   human chromosome 16p and 16q.";
```

```
RL   Genomics 60:295-308(1999).
RN   [2]
RP   SEQUENCE FROM N.A., AND REVISIONS TO 1263-1266.
RC   TISSUE=Brain;
RA   Nagase T., Kikuno R., Yamakawa H., Ohara O.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-1266 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99156230; PubMed=10048485;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 5:355-364(1998).
RN   [4]
RP   SEQUENCE OF 885-1972 FROM N.A.
RX   MEDLINE=93263189; PubMed=7684189;
RA   Matsuoka R., Yoshida M.C., Furutani Y., Imamura S., Kanda N.,
RA   Yanagisawa M., Masaki T., Takao A.;
RT   "Human smooth muscle myosin heavy chain gene mapped to chromosomal
RT   region 16q12.";
RL   Am. J. Med. Genet. 46:61-67(1993).
RN   [5]
RP   SEQUENCE OF 1093-1972 FROM N.A.
RC   TISSUE=Hippocampus;
RA   Okajima K.;
RL   Submitted (NOV-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Muscle contraction.
CC   -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC       heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC       and 2 regulatory light chain subunits (MLC-2).
CC   -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC   -!- TISSUE SPECIFICITY: Smooth muscle; expressed in the umbilical
CC       artery, bladder, esophagus and trachea.
CC   -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC       cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC       characteristic for alpha-helical coiled coils.
CC   -!- DISEASE: A chromosomal rearrangement, known as pericentric
CC       inversion Inv(16)(p13;q22), Produces a fusion protein that
CC       consists of the 165 N-terminal residues of CBF-beta (PEPB2) with
CC       the tail region of MYH11. This rearrangement is associated with
CC       acute myeloid leukemia of M4EO subtype.
CC   -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC       meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be
CC       split further into 2 globular subfragments (S1) and 1 rod-shaped
CC       subfragment (S2).
CC   -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC   -!- SIMILARITY: Contains 1 IQ domain.
CC   -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC       WWW="http://www.infobiogen.fr/services/chromcancer/Genes/MYH11ID43.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF001548; AAC31665.1; -.
DR   EMBL; U91323; AAC35212.1; -.
DR   EMBL; AB020673; BAA74889.2; ALT_INIT.
DR   EMBL; D10667; -; NOT_ANNOTATED_CDS.
DR   EMBL; X69292; CAA49154.1; -.
DR   HSSP; P10587; 1BR2.
DR   Genew; HGNC:7569; MYH11.
DR   MIM; 160745; -.
DR   GO; GO:0005859; C:muscle myosin; TAS.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR002017; Spectrin.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   Calmodulin-binding; ATP-binding; Methylation; Multigene family;
KW   Proto-oncogene; Chromosomal translocation.
FT   DOMAIN        1    785       MYOSIN HEAD-LIKE.
FT   DOMAIN      786    815       IQ.
FT   DOMAIN      844   1934       COILED COIL (POTENTIAL).
FT   DOMAIN     1935   1972       CARBOXYL-TERMINAL.
FT   NP_BIND     178    185       ATP (POTENTIAL).
FT   DOMAIN      661    683       ACTIN-BINDING (BY SIMILARITY).
FT   DOMAIN      762    776       ACTIN-BINDING (BY SIMILARITY).
FT   MOD_RES     129    129       METHYLATION (TRI-) (POTENTIAL).
FT   CONFLICT    887    889       EEK -> NSE (IN REF. 4).
FT   CONFLICT   1558   1558       T -> S (IN REF. 4).
FT   CONFLICT   1610   1611       KQ -> NE (IN REF. 4).
FT   CONFLICT   1786   1786       A -> S (IN REF. 5).
FT   CONFLICT   1958   1958       T -> L (IN REF. 4).
SQ   SEQUENCE   1972 AA;  227338 MW;  67665BB2AECE1277 CRC64;

Alignment Scores:
Pred. No.:                8.44e-09        Length:          1972
Score:                    319.00          Matches:         244
Percent Similarity:       36.82%          Conservative:    186
Best Local Similarity:    20.89%          Mismatches:      438
Query Match:              4.95%           Indels:          300
DB:                       1               Gaps:            49

US-09-602-362E-26 (1-3673) x MYHB_HUMAN (1-1972)

Qy        704 CCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTG---------GAATTTTTA 754
              |||||||||          ::::::      :::                  |||
Db        843 ProLeuLeuGlnValThrArgGlnGluGluGluMetGlnAlaLysGluAspGluLeuGln 862

Qy        755 CTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTT 814
                ||||||              |||   ||||||   |||
Db        863 LysThrLysGluArgGlnGlnLysAlaGluAsnGluLeuLysGluLeuGluGlnLysHis 882

Qy        815 GCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTC 874
              :::                 ::::::              ||||||          ::::::
Db        883 SerGlnLeuThrGluGluLysAsnLeuLeuGlnGluGlnLeuGlnAlaGluThrGluLeu 902

Qy        875 TTTGCTGAAGACATACATGGAATAACTGCAGAA-----------CGTTATGCTGCTGCT 922
              :::||||||                   ||||||              |||   ||||||
Db        903 TyrAlaGlu-----------------AlaGluGluMetArgValArgLeuAlaAlaLys 916

Qy        923 CGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAAT 982
              :::                 |||:::|||   |||   :::   ::::::
Db        917 Lys-----------------GlnGluLeuGluGluIleLeuHisGluMetGluAlaArg 930

Qy        983 CCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTG 1042
                :::               |||                        |||         :::
Db        931 LeuGluGluGluGluAspArgGlyGlnGlnLeuGlnAla------GluArgLysLysMet 948

Qy       1043 GCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCA 1102
              |||::::::     |||     |||   |||   |||          :::|||||||||
Db        949 AlaGlnGlnMetLeuAspLeuGluGluGlnLeuGluGlu---------GluGluAlaAla 965

Qy       1103 CGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAG 1162
              |||                 |||:::|||   |||   |||   |||:::
Db        966 Arg-----------------GlnLysLeuGln---LeuGluLysValThrAla------ 976

Qy       1163 TTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCT 1222
                 |||         ::::::         :::||||||
```


Db 977 ---GluAlaLysIleLysLysLeuGluAspGluIleLeu---------- 988
Qy 1223 GAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAA 1282
Db 989 --------------ValMetAspAspGlnAsnAsnLysLeuSer---------LysGlu 1000
Qy 1283 ACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAA 1342
Db 1001 ArgLysLeuLeuGluGluArgIleSerAspLeuThrThrAsnLeuAlaGluGluGluGlu 1020
Qy 1343 AAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACA 1402
Db 1021 Lys---AlaLysAsnLeuThrLysLeuLysAsnLysHisGluSerMetIleSerGluLeu 1039
Qy 1403 AATGTG---------- 1408
Db 1040 GluValArgLeuLysLysGluGluLysSerArgGlnGluLeuGluLysLeuLysArgLys 1059
Qy 1409 ---------GATGTGAGTTCTGTA---GAGCCTATATTCAGTCTTTTTGGCACACGGACT 1456
Db 1060 LeuGluGlyAspAlaSerAspPheHisGluGlnIleAlaAspLeuGlnAlaGlnIleAla 1079
Qy 1457 ATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGCTACCAAGATTATC 1516
Db 1080 GluLeuLysMetGlnLeuAlaLysLysGluGluLeuGlnAlaAlaLeuAlaArgLeu 1099
Qy 1517 TCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATC 1576
Db 1100 AspAspGluIleAlaGlnLysAsnAsnAlaLeu----------LysLysIle 1113
Qy 1577 AAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAATCCAAA--------- 1627
Db 1114 ArgGluLeuGluGlyHisIleSerAspLeuGlnGluAspLeuAspSerGluArgAlaAla 1133
Qy 1628 CGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAG 1687
Db 1134 ArgAsnLysAlaGluLysGlnLysArgAspLeuGlyGluGluLeuGlu---AlaLeuLys 1152
Qy 1688 ACTCAAGTG------TGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGA 1741
Db 1153 ThrGluLeuGluAspThrLeuAspSerThrAlaThrGlnGlnGluLeuArgAlaLysArg 1172
Qy 1742 GAA---------- 1744
Db 1173 GluGlnGluValThrValLeuLysLysAlaLeuAspGluGluThrArgSerHisGluAla 1192
Qy 1745 ----------GTAGAAGAGCTTCCTGAGAAGCCA 1768
Db 1193 GlnValGlnGluMetArgGlnLysHisAlaGlnAlaValGluGluLeuThrGluGlnLeu 1212
Qy 1769 TCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTG 1828
Db 1213 GluGlnPheLysArgAla----------LysAlaAsnLeuAsp 1223
Qy 1829 AAGAATGAACAAACATTG---------- 1846
Db 1224 LysAsnLysGlnThrLeuGluLysGluAsnAlaAspLeuAlaGlyGluLeuArgValLeu 1243
Qy 1847 ---AGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCT 1903
Db 1244 GlyGlnAlaLysGln------GluValGluHisLysLysLysLysLeuGluAlaGlnVal 1261
Qy 1904 TGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCT 1963
Db 1262 GlnGluLeuGlnSerLysCysSerAspGlyGluArgAlaArgAlaGluLeuAsnAspLys 1281
Qy 1964 ACACAT------CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTT 2017
Db 1282 ValHisLysLeuGlnAsnGluValGluSerValThrGlyMetLeuAsnGluAla------ 1299
Qy 2018 AAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTA 2077
Db 1300 ----------GluGlyLysAlaIle 1304
Qy 2078 GAATTA---AAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTG 2134
Db 1305 LysLeuAlaLysAspValAlaSerLeuSerSerGlnLeuGlnAspThrGlnGluLeuLeu 1324
Qy 2135 AAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGA 2194
Db 1325 GlnGluGluThrArgGlnLysLeuAsnValSerThrLysLeuArgGlnLeuGluGluGlu 1344
Qy 2195 GAAACACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGG 2254
Db 1345 ArgAsnSerLeuGlnAspGlnLeuAspGluGluMetGluAlaLysGlnAsnLeuGluArg 1364
Qy 2255 AAAGTTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCT 2314
Db 1365 HisIleSerThrLeuAsnIleGlnLeu---------- 1373
Qy 2315 CAGATGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAG 2374
Db 1374 ------------SerAspSerLysLysLys---------------LeuGlnAspPheAla 1384
Qy 2375 AGTTTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAA 2434
Db 1385 SerThrValGluAlaLeuGluGluGly--------------LysLysArgPheGlnLys 1399
Qy 2435 GAATTCGATACCTTAAGTGGAAAATTAGAAGAG----------TCT 2470
Db 1400 GluIleGluAsnLeuThrGlnGlnTyrGluGluLysAlaAlaAlaTyrAspLysLeuGlu 1419
Qy 2471 CCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAA 2530
Db 1420 LysThrLysAsnArgLeuGlnGlnGluLeuAspAspLeuValValAspLeuAspAsnGln 1439
Qy 2531 GCC---------- 2551
Db 1440 ArgGlnLeuValSerAsnLeuGluLysLysGlnArgLysPheAspGlnLeuLeuAlaGlu 1459
Qy 2552 ----------GAAACATTCAAAGCAGAGGATGTGAGTTCT 2581
Db 1460 GluLysAsnIleSerSerLysTyrAlaAspGluArgAspArgAlaGluAlaGluAlaArg 1479
Qy 2582 GTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAA 2641
Db 1480 GluLysGluThrLysAlaLeuSerLeuAlaArgAlaLeuGlu------GluAlaLeuGlu 1497
Qy 2642 GTTGAGGAAGACTTTAATCTTACTACCAAGGAG----------GGA 2677
Db 1498 AlaLysGluGluLeuGluArgThrAsnLysMetLeuLysAlaGluMetGluAspLeuVal 1517
Qy 2678 GCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCA 2737
Db 1518 SerSerLysAspAspValGlyLysAsnValHisGluLeuGluLysSerLysArgAlaLeu 1537
Qy 2738 CAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCA 2797
Db 1538 GluThrGlnMetGluGluMetLysThrGlnLeuGluGluLeuGluAspGluLeuGlnAla 1557
Qy 2798 ACTTCAGATTCTGAGATT---ATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCT 2854
Db 1558 ThrGluAspAlaLysLeuArgLeuGluVal------------AsnMetGlnAlaLeu--- 1572
Qy 2855 GAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCT--- 2911
Db 1573 LysGlyGlnPheGluArgAspLeuGlnAlaArgAspGluGlnAsnGluGluLysArgArg 1592
Qy 2912 ------GAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAAT 2965
Db 1593 GlnLeuGlnArgGlnLeuHis---GluTyrGluThrGluLeuGluAspGluArgLysGln 1611
Qy 2966 AAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAA 3025
Db 1612 ArgAlaLeuAlaAlaAlaAlaLysLysLysLeuGluGlyAspLeuLysAspLeuGluLeu 1631

```
Qy       3026 ATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAA 3085
                 |||          |||      :::|||||||||          :::      :::
Db       1632 GlnAlaAsp---------SerAlaIleLysGlyArgGlu---------GluAlaIleLys 1645

Qy       3086 CAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTG 3145
              |||:::                |||||||||:::                |||:::||||||
Db       1646 GlnLeuArgLysLeuGlnAlaGlnMetLysAsp------------PheGlnArgGluLeu 1661

Qy       3146 TCAGAAGCGAAAGAAATAAAATCACAGTTA---------------GAGAACCAAAAA--- 3187
                 :::|||:::     :::   ::::::               ||||||:::|||
Db       1662 GluAspAlaArgAlaSerArgAspGluIlePheAlaThrAlaLysGluAsnGluLysLys 1681

Qy       3188 ---GCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAG 3244
                    |||   :::|||      ::::::      |||      |||
Db       1682 AlaLysSerLeuGluAlaAspLeuMetGlnLeuGlnGluAspLeuAlaAlaAlaGluArg 1701

Qy       3245 AAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAA--------------- 3289
                 |||:::      |||:::   ||||||::::::      |||
Db       1702 AlaArgLysGlnAlaAspLeuGluLysGluGluLeuAlaGluGluLeuAlaSerSerLeu 1721

Qy       3290 ------------------GAGCAACTTAGGAAAAAGTTAGAAGTGAAA---CACCAACTT 3328
                               :::      :::::::::||||||   :::      ||||||
Db       1722 SerGlyArgAsnAlaLeuGlnAspGluLysArgArgLeuGluAlaArgIleAlaGlnLeu 1741

Qy       3329 GAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAG 3388
              |||:::   |||      :::         ::::::::::::::      :::   :::
Db       1742 GluGluGluLeuGluGluGluGlnGlyAsnMetGluAlaMetSerAspArgValArgLys 1761

Qy       3389 GTTTCTCACACTCATGAAAGT---GAAAATGATCTC------------------------ 3421
                 :::         |||      |||:::|||
Db       1762 AlaThrGlnGlnAlaGluGlnLeuSerAsnGluLeuAlaThrGluArgSerThrAlaGln 1781

Qy       3422 ------------------------------------------------------TTT 3424

Db       1782 LysAsnGluSerAlaArgGlnGlnLeuGluArgGlnAsnLysGluLeuArgSerLysLeu 1801

Qy       3425 CATGAA--------------AATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTG 3469
              ||||||                        |||      ||||||   |||:::
Db       1802 HisGluMetGluGlyAlaValLysSerLysPheLysSerThrIleAlaAlaLeuGluAla 1821

Qy       3470 GAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATT 3529
              ::::::|||    |||:::   |||   :::   :::      :::
Db       1822 LysIleAlaGlnLeuGluGluGlnValGluGlnGluAlaArgGluLysGlnAlaAlaThr 1841

Qy       3530 AAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACA 3589
              |||   |||::::::|||:::   :::|||:::      |||   |||:::  :::
Db       1842 LysSerLeuLysGlnLysAspLysLysLeuLysGluIleLeu---LeuGlnValGluAsp 1860

Qy       3590 GTAACAAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACG 3649
                 |||   |||   ||||||:::||||||   :::         |||
Db       1861 GluArgLysMetAlaGluGlnTyrLysGluGlnAlaGluLysGlyAsnAlaArgValLys 1880

Qy       3650 ATGCTGACTTCTAAATTGAAGGAA 3673
                    |||      :::|||:::|||
Db       1881 GlnLeuLysArgGlnLeuGluGlu 1888
```

```
RESULT 10
MYS2_DICDI
ID   MYS2_DICDI     STANDARD;      PRT;  2116 AA.
AC   P08799;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Myosin II heavy chain, non muscle.
GN   MHCA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87092266; PubMed=3540939;
RA   Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT   "Conserved protein domains in a myosin heavy chain gene from
RT   Dictyostelium discoideum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN   [2]
RP   PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC   STRAIN=AX2;
RX   MEDLINE=90353583; PubMed=2387408;
RA   Lueck-Vielmeter D., Schleicher M., Grabatin B., Wippler J.,
RA   Gerisch G.;
RT   "Replacement of threonine residues by serine and alanine in a
RT   phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL   FEBS Lett. 269:239-243(1990).
RN   [3]
RP   PHOSPHORYLATION SITES.
RX   MEDLINE=88112226; PubMed=2828113;
RA   Wagle G., Noegel A., Scheel J., Gerisch G.;
RT   "Phosphorylation of threonine residues on cloned fragments of the
RT   Dictyostelium myosin heavy chain.";
RL   FEBS Lett. 227:71-75(1988).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345066; PubMed=7619795;
RA   Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA   Rayment I.;
RT   "X-ray structures of the myosin motor domain of Dictyostelium
RT   discoideum complexed with MgADP.BeFx and MgADP.AlF4-.";
RL   Biochemistry 34:8960-8972(1995).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RX   MEDLINE=95345067; PubMed=7619796;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT   truncated head of Dictyostelium discoideum myosin to 2.7-A
RT   resolution.";
RL   Biochemistry 34:8973-8981(1995).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=96206189; PubMed=8611530;
RA   Smith C.A., Rayment I.;
RT   "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RT   Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL   Biochemistry 35:5404-5417(1996).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX   MEDLINE=97452580; PubMed=9305951;
RA   Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT   "X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes
RT   of the Dictyostelium discoideum myosin motor domain.";
RL   Biochemistry 36:11619-11628(1997).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX   MEDLINE=98070605; PubMed=9405148;
RA   Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT   "X-ray crystal structure and solution fluorescence characterization
RT   of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT   Dictyostelium discoideum myosin motor domain.";
RL   J. Mol. Biol. 274:394-407(1997).
CC   -!- FUNCTION: Myosin is a protein that binds to actin and has ATPase
CC       activity that is activated by actin.
CC   -!- SUBUNIT: Myosin II heavy chain is two-headed. It self-assembles
CC       into filaments. Hexamer of 2 heavy chain subunits (MHC), 2 alkali
CC       light chain subunits (MLC) and 2 regulatory light chain subunits
CC       (MLC-2).
CC   -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC       CORTEX.
CC   -!- DOMAIN: Each myosin heavy chain can be split into 1 light
CC       meromyosin (LMM) and 1 heavy meromyosin (HMM). It can be further
CC       split into 2 globular subfragments (S1) and 1 rod-shaped
CC       subfragment (S2).
CC   -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC       cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC       characteristic for alpha-helical coiled coils.
CC   -!- PTM: Phosphorylation inhibits thick filament formation and reduces
CC       the actin-activated ATPase activity.
```

```
CC   -!- MISCELLANEOUS: Dictyostelium myosin II has no K(2)EDTA ATPase
CC       activity, perhaps correlated with the absence of a Cys at the SH-1
CC       position (688).
CC   -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC   -!- SIMILARITY: Contains 1 IQ domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14628; AAA33227.1; -.
DR   PIR; A26655; A26655.
DR   PDB; 1MMA; 03-DEC-97.
DR   PDB; 1MMD; 17-AUG-96.
DR   PDB; 1MMG; 03-DEC-97.
DR   PDB; 1MMN; 03-DEC-97.
DR   PDB; 1MND; 17-AUG-96.
DR   PDB; 1MNE; 17-AUG-96.
DR   PDB; 1VOM; 23-DEC-96.
DR   PDB; 1LVK; 28-JAN-98.
DR   PDB; 1D0X; 20-DEC-00.
DR   PDB; 1D0Y; 20-DEC-00.
DR   PDB; 1D0Z; 20-DEC-00.
DR   PDB; 1D1A; 20-DEC-00.
DR   PDB; 1D1B; 20-DEC-00.
DR   PDB; 1D1C; 20-DEC-00.
DR   PDB; 1FMV; 20-DEC-00.
DR   PDB; 1FMW; 20-DEC-00.
DR   PDB; 1G8X; 17-JAN-01.
DR   PDB; 1JWY; 07-NOV-01.
DR   PDB; 1JX2; 07-NOV-01.
DR   DictyBase; DDB0002015; mhcA.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR008989; Myosin_S1_N.
DR   Pfam; PF00612; IQ; 2.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
KW   Calmodulin-binding; Methylation; Phosphorylation.
FT   DOMAIN        1    761       MYOSIN HEAD-LIKE.
FT   DOMAIN      762    791       IQ.
FT   DOMAIN      817   2116       COILED COIL (POTENTIAL).
FT   NP_BIND     179    186       ATP.
FT   DOMAIN      638    660       ACTIN-BINDING.
FT   DOMAIN      738    752       ACTIN-BINDING.
FT   MOD_RES     130    130       METHYLATION (DI-) (POTENTIAL).
FT   MOD_RES    1823   1823       PHOSPHORYLATION (BY MHCK).
FT   MOD_RES    1833   1833       PHOSPHORYLATION (BY MHCK).
FT   MOD_RES    2029   2029       PHOSPHORYLATION (BY MHCK).
FT   TURN          3      5
FT   TURN          7      8
FT   HELIX        10     15
FT   HELIX        25     28
FT   STRAND       34     37
FT   TURN         41     42
FT   STRAND       48     55
FT   STRAND       59     63
FT   STRAND       69     73
FT   HELIX        74     76
FT   STRAND       78     79
FT   HELIX        83     85
FT   TURN         86     87
FT   STRAND       90     90
FT   HELIX        91     93
FT   HELIX        99    110
FT   TURN        111    113
FT   STRAND      116    119
FT   TURN        120    121
FT   STRAND      122    126
FT   HELIX       137    142
FT   TURN        143    145
FT   HELIX       148    150
FT   HELIX       155    169
FT   TURN        170    170
FT   STRAND      173    179
FT   TURN        181    182
FT   HELIX       185    200
FT   HELIX       210    226
FT   STRAND      227    228
FT   TURN        234    235
FT   STRAND      236    237
FT   STRAND      240    247
FT   TURN        249    250
FT   STRAND      253    261
FT   HELIX       265    268
FT   TURN        269    269
FT   TURN        273    274
FT   STRAND      278    278
FT   HELIX       279    287
FT   HELIX       290    296
FT   TURN        297    297
FT   HELIX       301    303
FT   TURN        305    307
FT   TURN        316    317
FT   HELIX       320    334
FT   TURN        335    335
FT   HELIX       338    355
FT   TURN        356    356
FT   STRAND      360    360
FT   STRAND      368    368
FT   HELIX       373    382

Alignment Scores:
Pred. No.:                8.96e-09        Length:          2116
Score:                    318.50          Matches:         210
Percent Similarity:       37.96%          Conservative:    181
Best Local Similarity:    20.39%          Mismatches:      376
Query Match:              4.94%           Indels:          263
DB:                       1               Gaps:            44

US-09-602-362E-26 (1-3673) x MYS2_DICDI (1-2116)

Qy        857 CAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCT 916
              ::::::   |||::::::      ||||||                  ||||||      |||
Db       1201 LysLysLysValGluLeuAspLeuGluAspLysSerAlaGlnLeuAlaGluGluThrAla 1220

Qy        917 GCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTA--- 973
              |||   :::   ::::::   :::   :::::|||   :::   :::   ::::::
Db       1221 AlaLysGlnAlaLeuAspLysLeuLysLysLysLeuGluGlnGluLeuSerGluValGln 1240

Qy        974 -----------CCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCT------ACA 1015
                          |||   :::|||   |||   :::   |||:::
Db       1241 ThrGlnLeuSerGluAlaAsnAsnLysAsnValAsnSerAspSerThrAsnLysHisLeu 1260

Qy       1016 GGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTG 1075
                |||                  |||                  |||:::
Db       1261 GluThrSerPheAsnAsnLeuLysLeuGluLeuGluAlaGluGlnLysAlaLysGlnAla 1280

Qy       1076 CTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAA 1135
              |||||||||                  ||||||      |||      ::::::::::::
Db       1281 LeuGluLys--------------LysArgLeu------GlyLeuGluSerGluLeuLys 1293

Qy       1136 TGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAA 1195
                :::                  ||||||      ||||||
Db       1294 HisValAsn----------------------GluGlnLeuGluGluGlu------------ 1302
```

```
                     ::::::   :::      |||:::|||||||||              :::
Db       1887 ------GlyLysIleLysMetArgLeuGluLysGluLysArgAlaLeuGluGlyGluLeu 1904

Qy       3260 GATATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAA 3319
              :::   |||:::|||   :::         |||                  :::   |||
Db       1905 GluGluLeuArgGluThrValGluGluAlaGlu-----------------AspSerLys 1918

Qy       3320 CACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAAT 3379
                :::   ||||||:::   ||:::   ::::::||||||:::             |||
Db       1919 SerGluAlaGluGlnSerLysArgLeuValGluLeuGluLeuGluAspAlaArgArgAsn 1938

Qy       3380 TTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATG 3439
              |||   :::                |||        |||
Db       1939 LeuGlnLysGluIleAspAlaLysGluIleAlaGluAsp---------AlaLysSerAsn 1955

Qy       3440 TTGAAAAAGGAAATTGCC------------------------------------------ 3457
              |||::::::|||||
Db       1956 LeuGlnArgGluIleValGluAlaLysGlyArgLeuGluGluGluSerIleAlaArgThr 1975

Qy       3458 --------------------ATGCTAAAACTGGAAGTAGCCACACTGAAACATCAA--- 3493
                                   |||:::   |||:::      |||      |||
Db       1976 AsnSerAspArgSerArgLysArgLeuGluAlaGluIleAspAlaLeuThrAlaGlnVal 1995

Qy       3494 ------------------------CACCAGGTGAAGGAAAATAAATACTTTGAG---GAC 3526
                                        |||:::||||||||||||      |||   :::
Db       1996 AspAlaGluGlnLysAlaLysAsnGlnGlnIleLysGluAsnLysLysIleGluThrGlu 2015

Qy       3527 ATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAA 3586
              :::|||      ::::::|||      |||   :::   |||         :::
Db       2016 LeuLysGluTyrArgLysLysPheGlyGluSerGluLysThrLysThrLysGluPheLeu 2035

Qy       3587 ACAGTAACAAAAAGGGCATCTCAGTATAGA------------------------GAGCAG 3622
                |||   |||       :::   |||:::                       |||||
Db       2036 ValValGluLysLeuGluThrAspTyrLysArgAlaLysLysGluAlaAlaAspGluGln 2055

Qy       3623 CTTAAAGTTCTGACGGCAGAGAACACGATG 3652
                :::   ||||||   ||||||   :::
Db       2056 GlnGlnArgLeuThrValGluAsnAspLeu 2065
```

```
RESULT 11
MYHB_RABIT
ID   MYHB_RABIT     STANDARD;      PRT;  1972 AA.
AC   P35748;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Myosin heavy chain, smooth muscle isoform (SMMHC).
GN   MYH11.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92073350; PubMed=1961735;
RA   Babij P., Kelly C., Periasamy M.;
RT   "Characterization of a mammalian smooth muscle myosin heavy-chain
RT   gene: complete nucleotide and protein coding sequence and analysis of
RT   the 5' end of the gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10676-10680(1991).
CC   -!- FUNCTION: Muscle contraction.
CC   -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC       heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC       and 2 regulatory light chain subunits (MLC-2).
CC   -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC   -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC       cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC       characteristic for alpha-helical coiled coils.
CC   -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC       meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be
CC       split further into 2 globular subfragments (S1) and 1 rod-shaped
CC       subfragment (S2).
CC   -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC   -!- SIMILARITY: Contains 1 IQ domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M77812; AAA31395.1; -.
DR   HSSP; P10587; 1BR2.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR002017; Spectrin.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   Calmodulin-binding; ATP-binding; Methylation; Multigene family.
FT   DOMAIN        1    784       MYOSIN HEAD-LIKE.
FT   DOMAIN      785    807       IQ.
FT   DOMAIN      844   1934       COILED COIL (POTENTIAL).
FT   DOMAIN     1935   1972       CARBOXYL-TERMINAL.
FT   NP_BIND     178    185       ATP (POTENTIAL).
FT   DOMAIN      661    683       ACTIN-BINDING (BY SIMILARITY).
FT   DOMAIN      763    777       ACTIN-BINDING (BY SIMILARITY).
FT   MOD_RES     129    129       METHYLATION (TRI-) (POTENTIAL).
SQ   SEQUENCE   1972 AA;  227318 MW;  2061A224288D6A4C CRC64;
```

```
Alignment Scores:
Pred. No.:                1.33e-08        Length:          1972
Score:                    315.00          Matches:         227
Percent Similarity:       37.99%          Conservative:    203
Best Local Similarity:    20.05%          Mismatches:      460
Query Match:              4.89%           Indels:          242
DB:                       1               Gaps:            41

US-09-602-362E-26 (1-3673) x MYHB_RABIT (1-1972)

Qy        677 GAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAG 736
              |||   |||   ||||||              |||       :::|||::::::
Db        866 GluArgGlnGlnLysAlaGluSerGlu------LeuGlnGluLeuGlnGlnLysHisThr 883

Qy        737 CAAACTGTGGAA------TTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAG 790
              |||       |||         ||||||   :::         |||
Db        884 GlnLeuSerGluGluLysAsnLeuLeuGlnGluGlnLeuGlnAlaGluThrGluLeuTyr 903

Qy        791 TCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATG 850
              ::::::         :::   :::   :::                |||:::      :::
Db        904 AlaGluAlaGluGluMetArgValArgLeuAlaAlaLysLysGlnGluLeuGluGluIle 923

Qy        851 CTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGA------ATAACTGCA 904
              |||   :::         :::      |||:::      |||      :::   |||
Db        924 LeuHisGluMetGluAlaArgLeuGluGluGluGluAspArgGlyGlnGlnLeuGlnAla 943

Qy        905 GAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACAT--- 961
              ||||||         |||:::   :::      :::  :::||||||   |||
Db        944 GluArgLysLysMetAlaGlnGlnMetLeuAspLeuGluGluGlnLeuGluGluGluGlu 963

Qy        962 ------------------------------------------------ATACGAAAATTA 973
                                                               |||:::||||||
Db        964 AlaAlaArgGlnLysLeuGlnLeuGluLysValThrAlaGluAlaLysIleLysLysLeu 983

Qy        974 CCT--------------AAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGA 1018
```

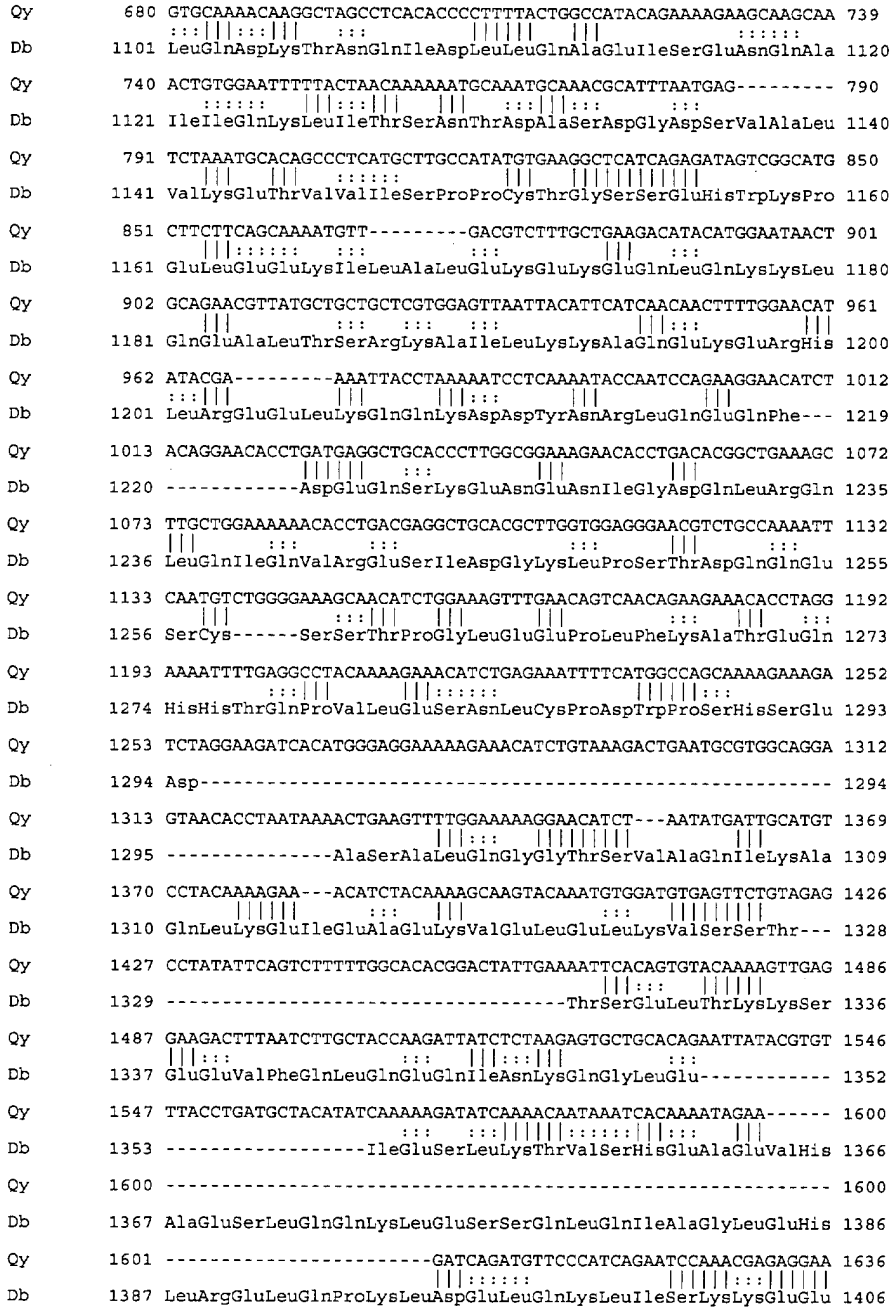

Db 984 GluAspAspIleLeuValMetAspAspGlnAsnAsnLys---------- 996
Qy 1019 ACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTG 1078
Db 997 LeuSerLysGluArgLysLeuLeuGluGluArgIleSerAspLeuThrThrAsnLeuAla 1016
Qy 1079 GAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGT 1138
Db 1017 GluGlu---GluGluLysAlaLysAsnLeuThrLys---------- 1027
Qy 1139 CTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATT 1198
Db 1028 ----------LeuLysAsnLysHisGluSerMetIleSerGluLeuGluValArgLeu 1043
Qy 1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
Db 1044 LysLysGluGluLysSerArgGlnGlu----------LeuGluLysLeuLysArg 1058
Qy 1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAG---ACTGAATGCGTGGCAGGAGTA 1315
Db 1059 LysMetAspGlyGluAlaSerAspLeuHisGluGlnIleAlaAspLeuGlnAlaGlnIle 1078
Qy 1316 ACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACA 1375
Db 1079 AlaGluLeuLysMetGlnLeuAlaLysLysGluGluGluLeuGlnAlaAlaLeuAlaArg 1098
Qy 1376 AAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTC 1435
Db 1099 LeuGluAspGluThrSerGlnLysAsnAsn---------- 1108
Qy 1436 AGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTT 1495
Db 1109 AlaLeuLysLysIleArgGluLeuGlu---GlyHisIleSerAspLeuGlnGluAspLeu 1127
Qy 1496 AATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGAT 1555
Db 1128 Asp----------SerGluArgAlaAlaArgAsnLysAlaGluLysGlnLys 1141
Qy 1556 GCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCA 1615
Db 1142 ArgAspLeuGlyGluGluLeuGluAlaLeuLysThrGluLeuGluAspThrLeuAspThr 1161
Qy 1616 TCAGAATCCAAACGAGAG----------GAAGATGAAGAATATTCTTGGGATTCTGGG 1663
Db 1162 ThrAlaThrGlnGlnGluLeuArgAlaLysArgGluGlnGluValThrValLeuLysLys 1181
Qy 1664 AGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAA 1723
Db 1182 AlaLeuAspGluGluThrArgSerHisGluAlaGlnValGlnGlu---MetArgGlnLys 1200
Qy 1724 GTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCT 1783
Db 1201 HisThrGlnVal---------ValGluGluLeuThrGluGlnLeuGluGlnPheLysArg 1217
Qy 1784 GCCGTN----------GAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAG 1831
Db 1218 AlaLysAlaAsnLeuAspLysThrLysGlnThrLeuGluLysGluAsnAlaAspLeuAla 1237
Qy 1832 AATGAACAAACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGAT 1891
Db 1238 GlyGluLeuArgValLeuGlyGlnAlaLysGlnGluValGluHisLysLysLysLysLeu 1257
Qy 1892 GAAGAAAATTCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTAT 1951
Db 1258 GluValGlnLeuGlnGluLeuGlnSerLysCysSerAspGlyGluArgAlaArgAlaGlu 1277
Qy 1952 TTACCCAAAGCTACACAT------CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAA 2005
Db 1278 LeuAsnAspLysValHisLysLeuGlnAsnGluValGluSerValThrGlyMetLeuSer 1297
Qy 2006 GAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCA 2065
Db 1298 GluAla----------Glu 1300
Qy 2066 AATAAAGCCTTAGAATTA---AAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAA 2122
Db 1301 GlyLysAlaIleLysLeuAlaLysGluValAlaSerLeuGlySerGlnLeuGlnAspThr 1320
Qy 2123 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2182
Db 1321 GlnGluLeuLeuGlnGluGluThrArgGlnLysLeuAsnValSerThrLysLeuArgGln 1340
Qy 2183 TTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCT 2242
Db 1341 LeuGluAspGluArgAsnSerLeuGlnGluGlnLeuAspGluGluMetGluAlaLysGln 1360
Qy 2243 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACA 2302
Db 1361 AsnLeuGluArgHisIleSerThrLeuAsnIleGlnLeu---------- 1373
Qy 2303 TTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCT 2362
Db 1374 ----------SerAspSerLysLysLys----------Leu 1380
Qy 2363 TGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCT 2422
Db 1381 GlnAspPheAlaSerThrValGluSerLeuGluGluGly----------LysLys 1395
Qy 2423 ACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG---TCTCCTGATAAA 2479
Db 1396 ArgPheGlnLysGluIleGluSerLeuThrGlnGlnTyrGluGluLysAlaAlaAlaTyr 1415
Qy 2480 GATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAA 2539
Db 1416 AspLysLeuGluLysThrLysAsnArgLeuGlnGlnGluLeuAspAspLeuValValAsp 1435
Qy 2540 TTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGT 2599
Db 1436 LeuAspAsnGlnArgGlnLeu----------ValSerAsnLeuGluLysLysGlnLys 1451
Qy 2600 CTTTTTGGCAAACCGACTACTGAAAATTCACAG---TCTACAAAAGTTGAGGAAGAC--- 2653
Db 1452 LysPheAspGlnLeuLeuAlaGluGluLysAsnIleSerSerLysTyrAlaAspGluArg 1471
Qy 2654 ------TTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAA 2707
Db 1472 AspArgAlaGluAlaGluAlaArgGluLysGluThrLysAlaLeuSerLeuAlaArgAla 1491
Qy 2708 CGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA 2767
Db 1492 LeuGluGluAlaLeuGluAlaLysGluGluLeuGluArgThrAsnLysMetLeuLysAla 1511
Qy 2768 GAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGT 2827
Db 1512 GluMet---------GluAspLeuValSerSerLysAspAspValGlyLysAsnValHis 1528
Qy 2828 GATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACA--- 2884
Db 1529 GluLeuGluLysSerLysArgAlaLeuGlu---ThrGlnMetGluGluMetLysThrGln 1547
Qy 2884 ---------- 2884
Db 1548 LeuGluGluLeuGluAspGluLeuGlnAlaThrGluAspAlaLysLeuArgLeuGluVal 1567
Qy 2885 ----------ACAAATGGCAAAATAGAAGAGTCT 2908
Db 1568 AsnMetGlnAlaLeuLysValGlnPheGluArgAspLeuGlnAlaArgAspGluGlnAsn 1587
Qy 2909 CCTGAAAAGCCTTCTCACTTT----------GAGCCTGCCACTGAAATGCAAAAC 2953
Db 1588 GluGluLysArgArgGlnLeuGlnArgGlnLeuHisGluTyrGluThrGluLeuGluAsp 1607
Qy 2954 TCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACT 3013
Db 1608 GluArgLysGlnArgAlaLeuAlaAlaAlaAlaLysLysLysLeuGluGlyAspLeuLys 1627

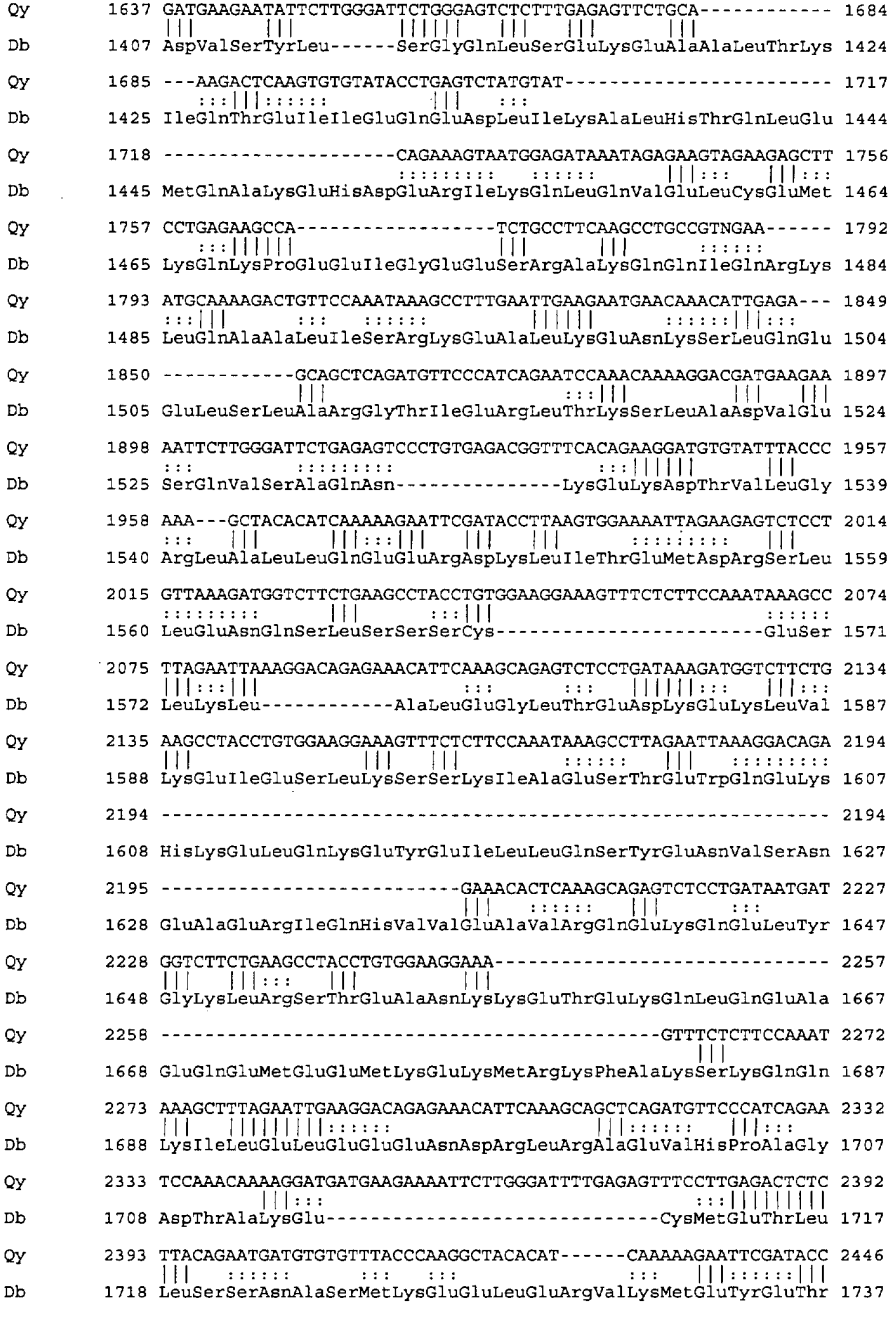
```
Qy       3014 ACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAA 3073
Db       1628 AspLeuGluLeuGlnAlaAsp---------SerAlaIleLysGlyArgGlu--------- 1641

Qy       3074 GATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTA 3133
Db       1642 GluAlaIleLysGlnLeuLeuLysLeuGlnAlaGlnMetLysAsp------------Phe 1657

Qy       3134 CAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTA---------------GAG 3178
Db       1658 GlnArgGluLeuGluAspAlaArgAlaSerArgAspGluIlePheAlaThrAlaLysGlu 1677

Qy       3179 AACCAAAAA------GCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAAT 3232
Db       1678 AsnGluLysLysAlaLysSerLeuGluAlaAspLeuMetGlnLeuGlnGluAspLeuAla 1697

Qy       3233 CAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAA--- 3289
Db       1698 AlaAlaGluArgAlaArgLysGlnAlaAspLeuGluLysGluGluLeuAlaGluGluLeu 1717

Qy       3290 ----------------------------GAGCAACTTAGGAAAAAGTTAGAAGTGAAA 3319
Db       1718 AlaSerSerLeuSerGlyArgAsnAlaLeuGlnAspGluLysArgArgLeuGluAlaArg 1737

Qy       3320 ---CACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTA------ 3370
Db       1738 IleAlaGlnLeuGluGluGluLeuGluGluGluGlnGlyAsnMetGluAlaMetSerAsp 1757

Qy       3371 --------------ACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAAT 3415
Db       1758 ArgValArgLysAlaThrGlnGlnAlaGluGlnLeuSerAsnGluLeuAlaThrGluArg 1777

Qy       3416 GATCTCTTTCATGAAAAT---------------------------------------- 3433
Db       1778 SerThrAlaGlnLysAsnGluSerAlaArgGlnGlnLeuGluArgGlnAsnLysGluLeu 1797

Qy       3434 ----------------------------------TGCATGTTGAAAAAGGAAATTGCC 3457
Db       1798 LysSerLysLeuGlnGluMetGluGlyAlaValLysSerLysPheLysSerThrIleAla 1817

Qy       3458 ATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATAC 3517
Db       1818 AlaLeuGluAlaLysIleAlaGlnLeuGluGluGlnValGluGlnGluAlaArgGluLys 1837

Qy       3518 TTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTG 3577
Db       1838 GlnAlaAlaAlaLysAlaLeuLysGlnArgAspLysLysLeuLysGluMetLeu---Leu 1856

Qy       3578 AAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACG 3637
Db       1857 GlnValGluAspGluArgLysMetAlaGluGlnTyrLysGluGlnAlaGluLysGlyAsn 1876

Qy       3638 GCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db       1877 AlaLysValLysGlnLeuLysArgGlnLeuGluGlu 1888
```

```
RESULT 12
GOB1_HUMAN
ID   GOB1_HUMAN     STANDARD;      PRT;  3259 AA.
AC   Q14789; Q14398;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Golgi autoantigen, golgin subfamily B member 1 (Giantin) (Macrogolgin)
DE   (Golgi complex-associated protein, 372-kDa) (GCP372).
GN   GOLGB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94187728; PubMed=7511208;
RA   Seelig H.P., Schranz P., Schroeter H., Wiemann C., Griffiths G.,
RA   Renz M.;
RT   "Molecular genetic analyses of a 376-kilodalton Golgi complex membrane
RT   protein (giantin).";
RL   Mol. Cell. Biol. 14:2564-2576(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94257116; PubMed=8198703;
RA   Seelig H.P., Schranz P., Schroeter H., Wiemann C., Griffiths G.,
RA   Renz M.;
RT   "Macrogolgin -- a new 376 kD Golgi complex outer membrane protein as
RT   target of antibodies in patients with rheumatic diseases and HIV
RT   infections.";
RL   J. Autoimmun. 7:67-91(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95100974; PubMed=7802676;
RA   Sohda M., Misumi Y., Fujiwara T., Nishioka M., Ikehara Y.;
RT   "Molecular cloning and sequence analysis of a human 372-kDa protein
RT   localized in the Golgi complex.";
RL   Biochem. Biophys. Res. Commun. 205:1399-1408(1994).
CC   -!- FUNCTION: May participate in forming intercisternal cross-bridges
CC       of the Golgi complex.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Golgi; membrane-associated.
CC   -!- DISEASE: Antigen in chronic rheumatoid arthritis and in the
CC       autoimmune disease Sjoegren's syndrome.
CC   -!- SIMILARITY: Belongs to the golgin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75304; CAA53052.1; -.
DR   EMBL; D25542; BAA05025.1; -.
DR   PIR; A56539; A56539.
DR   PIR; I52300; I52300.
DR   Genew; HGNC:4429; GOLGB1.
DR   MIM; 602500; -.
DR   GO; GO:0000139; C:Golgi membrane; TAS.
DR   GO; GO:0005795; C:Golgi stack; TAS.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0007030; P:Golgi organization and biogenesis; TAS.
KW   Golgi stack; Antigen; Coiled coil; Transmembrane.
FT   DOMAIN        1   3235       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   3236   3256       POTENTIAL.
FT   DOMAIN     3257   3259       LUMENAL (POTENTIAL).
FT   DOMAIN       48    593       COILED COIL (POTENTIAL).
FT   DOMAIN      677   1028       COILED COIL (POTENTIAL).
FT   DOMAIN     1062   1245       COILED COIL (POTENTIAL).
FT   DOMAIN     1301   1779       COILED COIL (POTENTIAL).
FT   DOMAIN     1828   3185       COILED COIL (POTENTIAL).
FT   DOMAIN     2420   2423       POLY-GLU.
FT   DOMAIN     2993   2996       POLY-SER.
FT   CONFLICT      1     39       MISSING (IN REF. 3).
FT   CONFLICT    215    215       A -> AQLSSM (IN REF. 3).
FT   CONFLICT   1765   1765       D -> G (IN REF. 3).
FT   CONFLICT   2950   2950       H -> D (IN REF. 3).
SQ   SEQUENCE   3259 AA;  376075 MW;  60376A20D8A178DD CRC64;
```

```
Alignment Scores:
Pred. No.:                2.13e-08        Length:          3259
Score:                    311.00          Matches:         235
Percent Similarity:       36.70%          Conservative:    212
Best Local Similarity:    19.29%          Mismatches:      425
Query Match:              4.82%           Indels:          346
DB:                       1               Gaps:            48
```

US-09-602-362E-26 (1-3673) x GOB1_HUMAN (1-3259)





```
Qy       2447 TTAAGTGGAAAATTAGAAGAG---TCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGT 2503
Db       1738 LeuSerLysLysPheGlnSerLeuMetSerGluLysAspSerLeuSerGluGluValGln 1757

Qy       2504 GGAATGAAAATTTCTCTTCCAAAT-----------------------AAAGCCTTAGAA 2539
Db       1758 AspLeuLysHisGlnIleGluAspAsnValSerLysGlnAlaAsnLeuGluAlaThrGlu 1777

Qy       2540 TTGAAGGACAGAGAAACATTCAAAGCAGAGGAT---------------------GTGAGT 2578
Db       1778 LysHisAspAsnGlnThrAsnValThrGluGluGlyThrGlnSerIleProGlyGluThr 1797

Qy       2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACT---ACTGAAAATTCACAGTCT 2635
Db       1798 GluGluGlnAspSerLeuSerMetSerThrArgProThrCysSerGluSerValProSer 1817

Qy       2636 ACAAAA---------------GTTGAGGAAGACTTTAATCTTACTACCAAG---GAGGGA 2677
Db       1818 AlaLysSerAlaAsnProAlaValSerLysAspPheSerSerHisAspGluIleAsnAsn 1837

Qy       2678 GCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCA 2737
Db       1838 TyrLeuGlnGlnIleAspGlnLeuLysGluArgIleAlaGlyLeuGluGlu--------- 1854

Qy       2738 CAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAA--- 2794
Db       1855 GluLysGlnLysAsnLysGluPheSerGlnThrLeuGluAsnGluLysAsnThrLeuLeu 1874

Qy       2795 ---------TCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAG 2845
Db       1875 SerGlnIleSerThrLysAspGlyGluLeuLysMetLeuGlnGluGluValThrLysMet 1894

Qy       2846 TGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAG 2905
Db       1895 AsnLeuLeuAsnGlnGlnIleGlnGluGluLeuSerArgValThr---LysLeuLysGlu 1913

Qy       2906 TCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAAT 2965
Db       1914 ThrAlaGlu---------------GluGluLysAspAspLeuGluGluArgLeuMetAsn 1928

Qy       2966 AAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAA 3025
Db       1929 Gln-----------------------------------------------LeuAlaGlu 1932

Qy       3026 ATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAA 3085
Db       1933 LeuAsnGlySerIle-----------------------------GlyAsnTyrCysGln 1942

Qy       3086 CAAATTACA---------GCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAA 3136
Db       1943 AspValThrAspAlaGlnIleLysAsnGluLeuLeuGluSerGluMetLysAsnLeuLys 1962

Qy       3137 AAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGG 3196
Db       1963 LysCysValSerGluLeuGluGluGluLysGlnGlnLeuValLysGluLysThrLysVal 1982

Qy       3197 GAACAAGAGCTCTGCAGTGTGAGATTG--------------------------CCTTTA 3229
Db       1983 GluSerGluIleArgLysGluTyrLeuGluLysIleGlnGlyAlaGlnLysGluProGly 2002

Qy       3230 AATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAA 3289
Db       2003 AsnLysSerHisAlaLysGluLeuGln---GluLeuLeuLysGluLysGlnGlnGluVal 2021

Qy       3290 GAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAA--------------CTTGAACAG 3334
Db       2022 LysGlnLeuGlnLysAspCys---IleArgTyrGlnGluLysIleSerAlaLeuGluArg 2040

Qy       3335 ACTCTCAGAATACAAGATATA--------------------GAATTGAAAAGTGTAACA 3373
Db       2041 ThrValLysAlaLeuGluPheValGlnThrGluSerGlnLysAspLeuGluIleThrLys 2060

Qy       3374 AGTAATTTGAATCAG---GTTTCTCACACTCATGAAAGTGAAAATGATCTC--------- 3421
Db       2061 GluAsnLeuAlaGlnAlaValGluHisArgLysLysAlaGlnAlaGluLeuAlaSerPhe 2080

Qy       3422 ----------------------------------------TTTCATGAAAATTGCATG 3439
Db       2081 LysValLeuLeuAspAspThrGlnSerGluAlaAlaArgValLeuAlaAspAsnLeuLys 2100

Qy       3440 TTGAAAAAGGAAATT----------------------GCCATGCTAAAACTGGAAGTA 3475
Db       2101 LeuLysLysGluLeuGlnSerAsnLysGluSerValLysSerGlnMetLysGlnLysAsp 2120

Qy       3476 GCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATT 3535
Db       2121 GluAspLeuGluArgArgLeuGluGlnAlaGluGluLysHisLeuLysGluLysLysAsn 2140

Qy       3536 TTACAAGAAAAG----------------------------------------AATGCT 3553
Db       2141 MetGlnGluLysLeuAspAlaLeuArgArgGluLysValHisLeuGluGluThrIleGly 2160

Qy       3554 GAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTAT 3613
Db       2161 GluIleGlnValThrLeuAsnLysLysAspLysGluVal------------GlnGlnLeu 2176

Qy       3614 AGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTG 3667
Db       2177 GlnGluAsnLeuAspSerThrValThrGlnLeuAlaAlaPheThrLysSerMet 2194
```

RESULT 13
GOA4_HUMAN
ID GOA4_HUMAN STANDARD; PRT; 2230 AA.
AC Q13439; Q13270; Q13654; Q14436;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230) (256
DE kDa golgin) (Golgin-245) (72.1 protein).
GN GOLGA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96215236; PubMed=8626529;
RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;
RT "Molecular characterization of trans-Golgi p230: a human peripheral
RT membrane protein encoded by a gene on chromosome 6p12-22 contains
RT extensive coiled-coil alpha-helical domains and a granin motif.";
RL J. Biol. Chem. 271:8328-8337(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Seelig H.P.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 131-2230 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96125112; PubMed=8537393;
RA Fritzler M.J., Lung C.-C., Hamel J.C., Griffith K.J., Chan E.K.L.;
RT "Molecular characterization of golgin-245, a novel Golgi complex
RT protein containing a granin signature.";
RL J. Biol. Chem. 270:31262-31268(1995).
RN [4]
RP SEQUENCE OF 524-672 FROM N.A.
RC TISSUE=Gastric fundus;
RA Balague C.;
RL Thesis (1994), Instituto municipal de investigacion medica, Spain.
CC -!- FUNCTION: May play a role in vesicular transport from the trans-
CC Golgi.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic, peripheral membrane protein
CC associated with the trans-Golgi network.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;

```
CC       Name=1;
CC         IsoId=Q13439-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q13439-2; Sequence=VSP_004272, VSP_004273;
CC       Name=3;
CC         IsoId=Q13439-3; Sequence=VSP_004274;
CC       Name=4;
CC         IsoId=Q13439-4; Sequence=VSP_004275;
CC   -!- DOMAIN: Extended rod-like protein with coiled-coil domains.
CC   -!- DISEASE: Antibodies against GOLGA4 are present in sera from
CC       patients with Sjoegren's syndrome (SS) [MIM:270150]. Sera from
CC       patients with SS often contain antibodies that react with normal
CC       components of the Golgi complex.
CC   -!- DISEASE: Antibodies against GOLGA4 are found in sera from
CC       hepatitis B patients.
CC   -!- SIMILARITY: Belongs to the golgin family.
CC   -!- SIMILARITY: Contains 1 GRIP domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U41740; AAC50434.1; -.
DR   EMBL; X82834; CAA58041.1; -.
DR   EMBL; U31906; AAC51791.1; -.
DR   EMBL; X76942; CAA54261.1; -.
DR   Genew; HGNC:4427; GOLGA4.
DR   MIM; 602509; -.
DR   MIM; 270150; -.
DR   GO; GO:0005802; C:Golgi trans face; TAS.
DR   GO; GO:0016192; P:vesicle-mediated transport; TAS.
DR   InterPro; IPR000237; GRIP_domain.
DR   Pfam; PF01465; GRIP; 1.
DR   PROSITE; PS50913; GRIP; 1.
KW   Antigen; Golgi stack; Coiled coil; Alternative splicing.
FT   DOMAIN      133   2185       COILED COIL (POTENTIAL).
FT   DOMAIN      252   2096       GLU-RICH.
FT   DOMAIN     2168   2215       GRIP.
FT   VARSPLIC   2154   2185       TPYKGGNLYHTDVSLFGEPTEPEYLRKVLFEY -> HLTKV
FT                                AICTIRMSHSLENLPNLSICEKCFLSI (in isoform
FT                                2).
FT                                /FTId=VSP_004272.
FT   VARSPLIC   2186   2230       Missing (in isoform 2).
FT                                /FTId=VSP_004273.
FT   VARSPLIC   2103   2109       Missing (in isoform 3).
FT                                /FTId=VSP_004274.
FT   VARSPLIC   2222   2230       FTSPRSGIF -> SWLRSSS (in isoform 4).
FT                                /FTId=VSP_004275.
FT   CONFLICT    188    188       R -> K (IN REF. 3).
FT   CONFLICT    220    220       Y -> H (IN REF. 3).
FT   CONFLICT    276    276       T -> A (IN REF. 3).
FT   CONFLICT    584    584       K -> E (IN REF. 3).
FT   CONFLICT    628    628       T -> A (IN REF. 3).
FT   CONFLICT    630    630       K -> E (IN REF. 3).
FT   CONFLICT    682    682       K -> N (IN REF. 3).
SQ   SEQUENCE   2230 AA;  261139 MW;  3BB733DB1EA86134 CRC64;

Alignment Scores:
Pred. No.:              2.35e-08      Length:         2230
Score:                  310.00        Matches:        189
Percent Similarity:     38.24%        Conservative:   159
Best Local Similarity:  20.77%        Mismatches:     352
Query Match:            4.81%         Indels:         210
DB:                     1             Gaps:           37

US-09-602-362E-26 (1-3673) x GOA4_HUMAN (1-2230)

Qy       1160 AAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATT---------TTGAGGCCTACA 1210
                 |||   :::|||          |||      :::::::::          |||   |||

Db          4 LysLeuLysGlnLysIleSerGluGluGlnGlnGlnLeuGlnGlnAlaLeuAlaProAla 23

Qy       1211 AAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATCACATGG 1270
                 :::   :::|||          |||   |||   :::   |||||||||
Db         24 GlnAlaSerSerAsnSer---SerThrProThrArgMetArgSerArg------------ 38

Qy       1271 GAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAATAAAACT 1330
                                                                 |||   :::   |||
Db         39 ----------------------------------------------ThrSerSerPheThr 43

Qy       1331 GAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAACATCTACA 1390
                 |||   |||::::::||||||   |||      :::      |||:::      :::
Db         44 GluGlnLeuAspGluGlyThrProAsnArgGluSerGlyAspThrGlnSerPheAlaGln 63

Qy       1391 AAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTC--------------- 1435
                 |||             :::   |||   |||||||||   :::|||
Db         64 Lys------LeuGlnLeuArgValProSerValGluSerLeuPheArgSerProIleLys 81

Qy       1436 ---AGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGAC 1492
                    |||||||||   :::   :::   |||:::      |||      |||
Db         82 GluSerLeuPheArgSerSerSerLysGluSerLeuValArgThrSerSerArgGluSer 101

Qy       1493 TTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCT 1552
                 |||                                                     |||
Db        102 LeuAsnArgLeuAsp-----------------------------------------LeuAsp 108

Qy       1553 GATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTC 1612
                    :::|||          |||
Db        109 SerSerThrAlaSerPheAsp---------------------------------------- 115

Qy       1613 CCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTT 1672
                 |||                  |||   |||:::         :::|||   ||||||
Db        116 ProProSerAspMetAspSerGluAlaGluAspLeuValGlyAsnSerAspSerLeuAsn 135

Qy       1673 GAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAG 1732
                 :::                        |||   :::   |||::::::
Db        136 Lys----------------------------GluGlnLeuIleGlnArgLeuArgArg 145

Qy       1733 ATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAA 1792
                 :::   |||   :::                |||   |||                |||   :::
Db        146 MetGluArgSerLeuSerSerTyrArgGlyLysTyrSerGluLeuValThrAlaTyrGln 165

Qy       1793 ATG------------CAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAA 1840
                 |||             :::|||   :::                :::   :::::::::
Db        166 MetLeuGlnArgGluLysLysLysLeuGlnGlyIleLeuSerGlnSerGlnAspLysSer 185

Qy       1841 ACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAAT 1900
                    |||   |||::::::          |||      |||   |||   :::
Db        186 LeuArgArgIleAlaGluLeu------ArgGluGluLeuGlnMetAspGlnGlnAlaLys 203

Qy       1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA 1960
                    |||             ::::::   :::||||||   |||:::
Db        204 LysHisLeuGlnGluGluPheAspAlaSerLeuGluGluLysAspGlnTyrIle------ 221

Qy       1961 GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAA 2020
                 :::      |||   :::         |||      :::|||         |||:::
Db        222 SerValLeuGlnThrGlnValSerLeuLeuLysGlnArgLeuArgAsnGlyProMetAsn 241

Qy       2021 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2080
                    :::|||||||||                |||||         |||||
Db        242 ValAspValLeuLysPro--------------------LeuProGln------LeuGlu 252

Qy       2081 TTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCT 2140
                    :::                  |||   |||:::|||:::   ||||||      :::|||
Db        253 ProGlnAlaGluValPheThrLysGluGluAsnProGluSerAspGly------GluPro 270

Qy       2141 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2200
                                    :::                  :::|||      ||||||
Db        271 --------------------ValValGluAspGlyThrSerValLysThrLeuGluThr 283

Qy       2201 CTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260
                 |||:::                  :::   |||||||||      |||         :::
Db        284 LeuGlnGlnArgValLysArgGlnGluAsnLeuLeuLys---ArgCysLysGluThrIle 302
```

```
Qy       2261 TCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320
Db        303 GlnSerHisLysGluGlnCysThrLeuLeuThrSerGlu---------------------- 315

Qy       2321 TTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAG------ 2374
Db        316 ------LysGluAlaLeuGlnGluGlnLeuAspGluArgLeuGlnGluLeuGluLysIle 333

Qy       2375 ------------------------------------------------------------AGT 2377
Db        334 LysAspLeuHisMetAlaGluLysThrLysLeuIleThrGlnLeuArgAspAlaLysAsn 353

Qy       2378 TTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAA 2437
Db        354 LeuIleGluGlnLeuGluGlnAspLysGlyMetValIleAlaGluThrLysArgGlnMet 373

Qy       2438 TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAG--- 2494
Db        374 HisGluThrLeuGluMetLysGluGluGluIleAlaGlnLeuArgSerArgIleLysGln 393

Qy       2495 ---CCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGA 2551
Db        394 MetThrThrGlnGlyGluGluLeu---------ArgGluGlnLysGluLysSerGluArg 410

Qy       2552 GAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAA 2611
Db        411 AlaAlaPhe------GluGluLeuGluLysAlaLeuSerThr------------AlaGln 424

Qy       2612 CCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACC--- 2668
Db        425 LysThrGluGluAlaArgArgLysLeuLysAlaGluMetAspGluGlnIleLysThrIle 444

Qy       2669 ---AAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGTGAT------------ 2713
Db        445 GluLysThrSerGluGluGluArgIleSerLeuGlnGlnGluLeuSerArgValLysGln 464

Qy       2714 ------ATTGGCATTATTGAACGAGCTCCACAAGATCAA---------ACAAATAAGATG 2758
Db        465 GluValValAspValMetLysLysSerSerGluGluGlnIleAlaLysLeuGlnLysLeu 484

Qy       2759 CCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATC 2818
Db        485 HisGluLysGluLeuAlaArgLysGlu----------------------GlnGluLeuThr 497

Qy       2819 TCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATA 2878
Db        498 LysLysLeuGlnThrArg-----------------GluArgGluPheGlnGluGlnMet 511

Qy       2879 AAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCC 2938
Db        512 LysValAla---------LeuGluLysSerGlnSerGluTyrLeuLysIleSerGlnGlu 528

Qy       2939 ACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTG 2998
Db        529 LysGluGlnGlnGluSerLeuAlaLeuGluGluLeuGluLeuGlnLysLysAlaIleLeu 548

Qy       2999 AGAGCAGATTCAACTACCCTA------------------------------TCAAAA 3025
Db        549 ThrGluSerGluAsnLysLeuArgAspLeuGlnGlnGluAlaGluThrTyrArgThrArg 568

Qy       3026 ATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAA 3085
Db        569 IleLeuGluLeuGluSerSerLeuGluLysSerLeuGlnGluAsnLysAsnGlnSerLys 588

Qy       3086 CAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTT--------------TGTGTA 3130
Db        589 AspLeuAlaValHisLeuGluAlaGluLysAsnLysHisAsnLysGluIleThrValMet 608

Qy       3131 CTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCT 3190
Db        609 ValGluLysHisLysThrGluLeuGluSerLeuLysHisGln---------GlnAspAla 625

Qy       3191 AAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAG--- 3247
Db        626 LeuTrpThrGluLysLeuGlnValLeuLysGlnGlnTyrGlnThrGluMetGluLysLeu 645

Qy       3248 AGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAG 3307
Db        646 ArgGluLysCysGluGlnGluLysGluThrLeuLeuLysAspLysGluIleIlePheGln 665

Qy       3308 TTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGATATA---------GAA 3358
Db        666 AlaHisIleGluGluMetAsnGluLysThrLeuGluLysLeuAspValLysGlnThrGlu 685

Qy       3359 TTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGAT 3418
Db        686 LeuGluSerLeuSerSerGluLeuSerGluValLeuLysAlaArgHisLys--------- 702

Qy       3419 CTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCC 3478
Db        703 --------------------LeuGluGluGluLeuSerValLeuLysAspGlnThrAsp 715

Qy       3479 ACACTGAAACATCAACACCAGGTGAAG-----------GAAAATAAATACTTTGAGGAC 3526
Db        716 LysMetLysGlnGluLeuGluAlaLysMetAspGluGlnLysAsnHisHisGlnGlnGln 735

Qy       3527 ATT---AAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACC--------------- 3568
Db        736 ValAspSerIleIleLysGluHisGluValSerIleGlnArgThrGluLysAlaLeuLys 755

Qy       3569 --------------CTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTAT 3613
Db        756 AspGlnIleAsnGlnLeuGluLeuLeuLeuLysGluArgAspLysHisLeuLysGluHis 775

Qy       3614 AGAGAGCAGCTTAAAGTTCTGACGGCAGAG 3643
Db        776 GlnAlaHisValGluAsnLeuGluAlaAsp 785
```

RESULT 14
MYH9_RAT
ID MYH9_RAT STANDARD; PRT; 1961 AA.
AC Q62812;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, nonmuscle type A (Cellular myosin heavy chain,
DE type A) (Nonmuscle myosin heavy chain-A) (NMMHC-A).
GN MYH9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi O.H., Park C., Itoh K., Adelstein R.S., Beaven M.A.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Cellular myosin appears to play a role in cytokinesis,
CC cell shape, and specialized functions such as secretion and
CC capping.
CC -!- SUBUNIT: Myosin is a hexameric protein that consists of 2 heavy
CC chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2
CC regulatory light chain subunits (MLC-2).
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC --

```
CC   ---------------------------------------------------------------------------
DR   EMBL; U31463; AAA74950.1; -.
DR   HSSP; P10587; 1BR2.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW   Coiled coil; Multigene family.
FT   DOMAIN        1    778       MYOSIN HEAD-LIKE.
FT   DOMAIN      779    808       IQ.
FT   DOMAIN      841   1927       COILED COIL (POTENTIAL).
FT   NP_BIND     174    181       ATP (POTENTIAL).
FT   DOMAIN      654    676       ACTIN-BINDING.
SQ   SEQUENCE   1961 AA;  226336 MW;  9B9876D9681FB19E CRC64;

Alignment Scores:
Pred. No.:                2.77e-08        Length:          1961
Score:                    308.50          Matches:         227
Percent Similarity:       38.36%          Conservative:    198
Best Local Similarity:    20.49%          Mismatches:      454
Query Match:              4.78%           Indels:          229
DB:                       1               Gaps:            43

US-09-602-362E-26 (1-3673) x MYH9_RAT (1-1961)

Qy        704 CCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAA 763
              |||||||||   :::|||:::        :::                |||||   |||
Db        836 ProLeuLeuAsnSerIleArgHisGluAspGlu--------------LeuLeuAlaLys 850

Qy        764 AATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCC------------------ 805
                  |||                |||           |||
Db        851 GluAlaGluLeuThrLysValArgGluLysHisLeuAlaAlaGluAsnArgLeuThrGlu 870

Qy        806 ---------------------CTCATG--------------------------------- 811
                                   ||||||
Db        871 MetGluThrMetGlnSerGlnLeuMetAlaGluLysLeuGlnLeuGlnGluGlnLeuGln 890

Qy        812 ------CTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTT------CAG 859
                       :::|||      :::   |||:::          |||         :::
Db        891 AlaLysThrGluLeuCysAlaGluAlaGluGluLeuArgAlaArgLeuThrAlaLysLys 910

Qy        860 CAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCTGCT 919
              |||   ::::::           |||:::                |||
Db        911 GlnGluLeuGluGluIleCysHisAspLeuGluAlaArgValGluGluGluGluArg 930

Qy        920 GCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAA 979
                  :::   :::           :::::::::   ::::::|||::::::|||   :::
Db        931 CysGlnTyrLeuGlnAlaGluLysLysLysMetGlnGlnAsnIleGlnGluLeuGluGlu 950

Qy        980 AATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCC 1039
                  :::                                    :::   |||
Db        951 GlnLeuGluGlu-----------------------------GluGluSerAlaArgGln 960

Qy       1040 TTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCT 1099
                  :::             |||   |||:::   |||:::|||      :::|||
Db        961 LysLeuGlnLeuGluLysValThrThrGluAlaLysLeuLysLysLeu---GluGluAsp 979

Qy       1100 GCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGA 1159
                  ::::::|||      :::   |||:::          |||
Db        980 GlnIleIleMetGluAspGlnAsnCysLysLeuAlaLysGluLysLysLeuGluAsp 999

Qy       1160 AAGTTTGAACAGTCAACA--------------GAAGAAACACCTAGGAAAATTTTGAGG 1204
                  :::      :::   |||               ||||||         |||   |||

Db       1000 ArgValAlaGluPheThrThrAspLeuMetGluGluGluGluLysSerLysSerLeuAla 1019

Qy       1205 CCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAAGATC 1264
                  |||           |||                  :::||||||   |||:::
Db       1020 LysLeuLysAsnLysHisGluAlaMetIleThrAspLeuGluGluArgLeuArgArg--- 1038

Qy       1265 ACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAAT 1324
                  |||||||||:::           :::                |||
Db       1039 ------GluGluLysGlnArgGlnGluLeuGluLys--------------ThrArgArg 1051

Qy       1325 AAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCA------------------ 1366
                  |||   |||                  |||:::   ||||||
Db       1052 LysLeuGluGlyAspSerThrAspLeuSerAspGlnIleAlaGluLeuGlnAlaGlnIle 1071

Qy       1367 --------------------TGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAAT 1405
                                    :::|||          |||
Db       1072 AlaGluLeuLysMetGlnLeuAlaLysLysGluGluGluLeuGlnAlaAlaLeuAlaArg 1091

Qy       1406 GTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAAT 1465
                  |||:::      :::   :::         :::|||          |||   :::|||
Db       1092 ValGluGluGluAlaAlaGlnLysAsnMetAlaLeuLysLysIleArgGluLeuGlu--- 1110

Qy       1466 TCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGT 1525
                  :::|||   :::::::::::::||||||                    |||:::
Db       1111 ThrGlnIleSerGluLeuGlnGluAspLeuGlu-----------------SerGluArg 1124

Qy       1526 GCTGCACAGAATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATA 1585
                  |||   :::|||                          ::::::::::::   :::
Db       1125 AlaCysArgAsnLysAlaGluLysGlnLysArgAspLeuGlyGluGluLeuGluAlaLeu 1144

Qy       1586 AATCACAAAATAGAAGATCAGATGTTCCCATCAGAATCCAAACGAGAG------------ 1633
                  ::::::||||||   :::       :::   :::::::::|||
Db       1145 LysThrGluLeuGluAspThrLeuAspSerThrAlaAlaGlnGlnGluLeuArgSerLys 1164

Qy       1634 GAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAA 1693
                  ::::::|||   |||           :::|||   |||       |||||||||
Db       1165 ArgGluGlnGluValSerIleLeuLysLysThrLeu---GluAspGluAlaLysThrHis 1183

Qy       1694 GTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGAAGTAGAAGAG 1753
                  :::   |||   ||||||   :::       |||:::   |||
Db       1184 GluAlaGlnIleGlnGluMetArgGlnLysHisSerGlnAlaValGluGluLeuAlaGlu 1203

Qy       1754 CTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAG---ACTGTTCCA 1810
                  |||:::          |||      :::|||      |||   |||:::
Db       1204 GlnLeuGluGlnThrLysArgValLysAlaThrLeuGluLysAlaLysGlnThrLeuGlu 1223

Qy       1811 AATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATGTTCCCATCA 1870
                  |||:::      ||||||   ||||||                        |||
Db       1224 AsnGluArgGlyGluLeuAlaAsnGluValLysAlaLeuLeuGlnGlyLysGlyAspSer 1243

Qy       1871 GAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAG---------AGTCCC 1921
                  |||   |||:::|||       |||             :::   :::         |||
Db       1244 GluHisLysArgLysLysValGluAlaGlnLeuGlnGluLeuGlnValLysPheSerGlu 1263

Qy       1922 TGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTC 1981
                  |||   |||   :::   ::::::          :::      |||   |||
Db       1264 GlyGluArgVal---ArgThrGluLeuAlaAspLysValSerLysLeuGlnValGluLeu 1282

Qy       1982 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACC 2041
                  |||:::::::::|||   |||   :::|||      |||         |||
Db       1283 AspSerValThrGlyLeuLeuAsnGlnSerAspSerLysSerSerLysLeu--------- 1299

Qy       2042 TGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTC 2101
                                                      ||||||
Db       1300 ------------------------------------ThrLysAspPheSerAlaLeu 1306

Qy       2102 AAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT 2161
                  :::::::::      |||           ||||||::              :::|||:::|||
Db       1307 GluSerGlnLeuGlnAspThrGlnGluLeuLeuGlnGluGluAsnArgGlnLysLeuSer 1326

Qy       2162 CTTCCAAATAAAGCCTTAGAATTAAAGGAC--------------AGAGAAACACTCAAA 2206
                  |||      |||          :::::::::|||                ||||||   |||:::
Db       1327 LeuSerThrLysLeuLysGlnMetGluAspGluLysAsnSerPheArgGluGlnLeuGlu 1346
```

```
Qy       2207 GCAGAGTCTCCTGATAATGATGGTCTTCTGAAG---------CCTACCTGTGGAAGGAAA 2257
Db       1347 GluGluGluGluGluAlaLysArgAsnLeuGluLysGlnIleAlaThrLeuHisAlaGln 1366

Qy       2258 GTTTCTCTTCCAAATAAAGCTTTAGAA--------------TTGAAGGACAGAGAAACA 2302
Db       1367 ValThrAspMetLysLysLysMetGluAspGlyValGlyCysLeuGluThrAlaGluGlu 1386

Qy       2303 TTCAAAGCAGCTCAGATGTTCCCATCAGAA------TCCAAACAAAAGGATGATGAAGAA 2356
Db       1387 AlaLysArgArgLeuGlnLysAspLeuGluGlyLeuSerGlnArgLeuGluGluLysVal 1406

Qy       2357 AATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAAT------GATGTGTGT 2410
Db       1407 AlaAlaTyrAspLysLeuGluLysThrLysThrArgLeuGlnGlnGluLeuAspAspLeu 1426

Qy       2411 TTACCCAAGGCTACACAT--------------------------------CAAAAAGAA 2437
Db       1427 LeuValAspLeuAspHisGlnArgGlnSerValSerAsnLeuGluLysLysGlnLysLys 1446

Qy       2438 TTCGAT--------------------ACCTTAAGTGGAAAATTAGAAGAGTCTCCTGAT 2476
Db       1447 PheAspGlnLeuLeuAlaGluGluLysThrIleSerAlaLysTyrAlaGluGluArgAsp 1466

Qy       2477 AAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTA 2536
Db       1467 Arg----------------------------------------------AlaGluAla 1470

Qy       2537 GAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTC 2596
Db       1471 GluAlaArgGluLysGluThrLysAlaLeuSerLeuAlaArgAlaLeuGluGluAlaMet 1490

Qy       2597 AGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTT 2656
Db       1491 GluGlnLysAlaGluLeuGluArgLeuAsnLysGlnPheArgThrGluMetGluAspLeu 1510

Qy       2657 AATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGTGATATT 2716
Db       1511 ---MetSerSerLysAspAspValGlyLysSerVal--------------HisGluLeu 1524

Qy       2717 GGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGA 2776
Db       1525 GluLysSerAsnArgAlaLeuGluGlnGlnValGluGluMetLysThrGlnLeuGluGlu 1544

Qy       2777 AGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATT---ATCTCTGTGAGTGATACA 2833
Db       1545 LeuGluAspGluLeuGlnAlaThrGluAspAlaLysLeuArgLeuGluVal--------- 1561

Qy       2834 CAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGC 2893
Db       1562 ---AsnLeuGlnAlaMet---LysAlaGlnPheGluArgAspLeuGln---------Gly 1576

Qy       2894 AAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTT--------------GAGCCTGCC 2938
Db       1577 ArgAspGluGlnSerGluGluLysLysLysGlnLeuValArgGlnValArgGluMetGlu 1596

Qy       2939 ACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTG 2998
Db       1597 AlaGluLeuGluAspGluArgLysGlnArgSerIleAlaMetAlaAlaArgLysLysLeu 1616

Qy       2999 AGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCA---------------------- 3037
Db       1617 GluMetAspLeuLysAspLeuGluAlaHisIleAspThrAlaAsnLysAsnArgGluGlu 1636

Qy       3038 --------------------------------CTTCCTTCTTGTGAAAGAGGAAGGGAA 3064
Db       1637 AlaIleLysGlnLeuArgLysLeuGlnAlaGlnMetLysAspCysMetArgAspValAsp 1656

Qy       3065 CTTAAAAAAGATAACTGTGAACAAATTACAGCA---------------------AAAATG 3103
Db       1657 AspThrArgAlaSerArgGluGluIleLeuAlaGlnAlaLysGluAsnGluLysLysLeu 1676

Qy       3104 GAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATA 3163
Db       1677 LysSerMetGluAlaGluMetIleGlnLeuGlnGluGluLeuAlaAlaAlaGluArgAla 1696

Qy       3164 AAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTG------ 3217
Db       1697 LysArgGlnAlaGlnGlnGluArgAspGluLeuAlaAspGluIleAlaAsnSerSerGly 1716

Qy       3218 ------AGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGA---------AATGTCGAT 3262
Db       1717 LysGlyAlaLeuAlaLeu---------GluGluLysArgArgLeuGluAlaLeuIleAla 1733

Qy       3263 ATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACAC 3322
Db       1734 LeuLeuGluGluGluLeuGlu---GluGluGlnGlyAsnThrGluLeu------------ 1748

Qy       3323 CAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTG 3382
Db       1749 ---IleAsnAspArgLeuLysLysAlaAsnLeuGlnIleAspGlnIleAsnThrAspLeu 1767

Qy       3383 AAT---CAGGTTTCTCACACTCATGAAAGTGAAAAT---------GATCTCTTTCATGAA 3430
Db       1768 AsnLeuGluArgSerHisAlaGlnLysAsnGluAsnAlaArgGlnGlnLeuGluArgGln 1787

Qy       3431 AATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACAT 3490
Db       1788 AsnLysGluLeuLys----------------------------AlaLysLeuGlnGlu 1797

Qy       3491 CAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAAT 3550
Db       1798 MetGluSerAlaValLysSerLysTyrLysAlaSerIleAlaAlaLeuGluAlaLysIle 1817

Qy       3551 GCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAG 3610
Db       1818 AlaGlnLeuGluGluGlnLeuAspAsnGluThrLysGluArgGlnAlaAlaSerLysGln 1837

Qy       3611 TATAGAGAGCAGCTTAAAGTTCTG 3634
Db       1838 ValArgArgAlaGluLysLysLeu 1845
```

```
RESULT 15
MYHB_CHICK
ID   MYHB_CHICK     STANDARD;      PRT;  1978 AA.
AC   P10587;
DT   01-JUL-1989 (Rel. 11, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Myosin heavy chain, gizzard smooth muscle.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88118918; PubMed=2892941;
RA   Yanagisawa M., Hamada Y., Katsuragawa Y., Imamura M., Mikawa T.,
RA   Masaki T.;
RT   "Complete primary structure of vertebrate smooth muscle myosin heavy
RT   chain deduced from its complementary DNA sequence. Implications on
RT   topography and function of myosin.";
RL   J. Mol. Biol. 198:143-157(1987).
RN   [2]
RP   REVISIONS.
RA   Masaki T.;
RL   Submitted (FEB-1989) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-203.
RX   MEDLINE=88032919; PubMed=3312184;
RA   Maita T., Onishi H., Yajima E., Matsuda G.;
RT   "Amino acid sequence of the amino-terminal 24 kDa fragment of the
RT   heavy chain of chicken gizzard myosin.";
RL   J. Biochem. 102:133-145(1987).
```

```
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 1-818, AND REVISIONS 204-215.
RX   MEDLINE=98412652; PubMed=9741621;
RA   Dominguez R., Freyzon Y., Trybus K.M., Cohen C.;
RT   "Crystal structure of a vertebrate smooth muscle myosin motor domain
RT   and its complex with the essential light chain: visualization of the
RT   pre-power stroke state.";
RL   Cell 94:559-571(1998).
CC   -!- FUNCTION: Muscle contraction.
CC   -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC       heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC       and 2 regulatory light chain subunits (MLC-2).
CC   -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC   -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC       cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC       characteristic for alpha-helical coiled coils.
CC   -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC       meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be
CC       split further into 2 globular subfragments (S1) and 1 rod-shaped
CC       subfragment (S2).
CC   -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC   -!- SIMILARITY: Contains 1 IQ domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X06546; CAA29793.1; -.
DR   PIR; S03166; S03166.
DR   PDB; 1BR1; 09-SEP-98.
DR   PDB; 1BR2; 09-SEP-98.
DR   PDB; 1BR4; 09-SEP-98.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001609; myosin_head.
DR   InterPro; IPR004009; Myosin_N.
DR   InterPro; IPR002928; Myosin_tail.
DR   InterPro; IPR002017; Spectrin.
DR   Pfam; PF00612; IQ; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   Pfam; PF02736; Myosin_N; 1.
DR   Pfam; PF01576; Myosin_tail; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   ProDom; PD000355; myosin_head; 1.
DR   SMART; SM00015; IQ; 1.
DR   SMART; SM00242; MYSc; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW   ATP-binding; Methylation; Calmodulin-binding; Multigene family;
KW   3D-structure.
FT   INIT_MET      0      0
FT   DOMAIN        1    790       MYOSIN HEAD-LIKE.
FT   DOMAIN      791    820       IQ.
FT   DOMAIN      849   1978       RODLIKE TAIL (S2 AND LMM DOMAINS).
FT   DOMAIN      849   1978       COILED COIL (POTENTIAL).
FT   NP_BIND     176    183       ATP.
FT   DOMAIN      666    688       ACTIN-BINDING.
FT   DOMAIN      767    781       ACTIN-BINDING.
FT   MOD_RES       1      1       BLOCKED.
FT   MOD_RES     127    127       METHYLATION (TRI-) (POTENTIAL).
FT   CONFLICT    127    127       MISSING (IN REF. 3).
FT   CONFLICT    204    215       KDTSITQGPSFS -> RTPASLKVHLFP (IN REF. 1).
FT   STRAND       34     37
FT   STRAND       43     47
FT   STRAND       56     60
FT   STRAND       66     69
FT   TURN         72     73
FT   STRAND       75     76
FT   HELIX        80     82
FT   TURN         83     84
FT   STRAND       87     87
FT   TURN         88     89
FT   HELIX        96    108
FT   TURN        109    110
FT   STRAND      113    115
FT   STRAND      120    123
FT   TURN        130    131
FT   HELIX       134    139
FT   TURN        140    142
FT   HELIX       145    147
FT   HELIX       152    166
FT   TURN        167    167
FT   STRAND      170    175
FT   TURN        178    179
FT   HELIX       182    197
FT   HELIX       218    223
FT   TURN        224    225
FT   HELIX       226    234
FT   STRAND      235    236
FT   TURN        242    243
FT   STRAND      244    245
FT   STRAND      248    255
FT   TURN        257    258
FT   STRAND      261    269
FT   HELIX       273    276
FT   TURN        277    277
FT   TURN        281    282
FT   STRAND      286    286
FT   HELIX       287    295
FT   HELIX       298    301
FT   TURN        302    305
FT   TURN        309    311
FT   TURN        313    314
FT   TURN        316    317
FT   HELIX       327    341
FT   TURN        342    342
FT   HELIX       345    363
FT   STRAND      367    367
FT   STRAND      376    376
FT   HELIX       381    390
FT   TURN        391    391
FT   HELIX       394    402
FT   HELIX       419    449
FT   STRAND      458    464
FT   STRAND      474    474
FT   HELIX       476    506
FT   TURN        507    508
FT   HELIX       521    528
FT   HELIX       536    543
FT   HELIX       551    561
FT   TURN        562    564
FT   TURN        566    567
FT   STRAND      568    570
FT   STRAND      579    584
FT   TURN        585    586
FT   STRAND      587    592
FT   TURN        594    595
FT   HELIX       596    601
FT   HELIX       606    613
FT   TURN        614    614
FT   HELIX       618    623
FT   TURN        624    624
FT   HELIX       658    674
FT   TURN        675    675
FT   STRAND      677    684
FT   TURN        692    693
FT   HELIX       697    706
FT   TURN        707    708
FT   HELIX       709    718
FT   STRAND      722    725
FT   HELIX       726    733
FT   HELIX       734    737
FT   TURN        738    739
```

```
FT   HELIX        748    759
FT   TURN         760    760
FT   HELIX        763    765
FT   STRAND       766    768
FT   STRAND       772    775
FT   TURN         777    778
FT   HELIX        779    787
SQ   SEQUENCE   1978 AA;  228663 MW;  B7B6C923E5273D93 CRC64;

Alignment Scores:
Pred. No.:              7.25e-08        Length:         1978
Score:                  300.00          Matches:        228
Percent Similarity:     38.46%          Conservative:   202
Best Local Similarity:  20.39%          Mismatches:     424
Query Match:            4.65%           Indels:         264
DB:                     1               Gaps:           49

US-09-602-362E-26 (1-3673) x MYHB_CHICK (1-1978)

Qy        788 GAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGC 847
Db        884 GluGlnLysHisThrGlnLeu-----------CysGlu---GluLysAsnLeuLeuGln 898

Qy        848 ATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAA 907
Db        899 GluLysLeuGlnAlaGluThrGluLeuTyrAlaGlu-----------------AlaGlu 912

Qy        908 ------------CGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTG 955
Db        913 GluMetArgValArgLeuAlaAlaLysLys-----------------GlnGluLeuGlu 926

Qy        956 GAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACA 1015
Db        927 GluIleLeuHisGluMetGluAlaArgIleGlu------GluGluGluGluArgSerGln 944

Qy       1016 GGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTG 1075
Db        945 GlnLeuGlnAlaGluLysLysLysMetGlnGlnGlnMetLeuAspLeuGluGluGlnLeu 964

Qy       1076 CTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAA 1135
Db        965 GluGlu---------GluGluAlaAlaArg-----------------GlnLysLeuGln 975

Qy       1136 TGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAA 1195
Db        976 LeuGluLysValThrAlaAspGlyLysIleLysLysMetGluAspAsp------------ 991

Qy       1196 ATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCT 1255
Db        992 IleLeu-----------------------------------IleMetGluAspGlnAsn 999

Qy       1256 AGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTA 1315
Db       1000 AsnLysLeuThr---------LysGluArgLysLeuLeuGluGluArgValSerAspLeu 1016

Qy       1316 ACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACA 1375
Db       1017 ThrThrAsnLeuAlaGluGluGluGluLys---AlaLysAsnLeuThrLysLeuLysAsn 1035

Qy       1376 AAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTA------------ 1423
Db       1036 LysHisGluSerMetIleSerGluLeuGluValArgLeuLysLysGluGluLysSerArg 1055

Qy       1424 ------------------------------------------------------GAGCCT 1429
Db       1056 GlnGluLeuGluLysIleLysArgLysLeuGluGlyGluSerSerAspLeuHisGluGln 1075

Qy       1430 ATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAA 1489
Db       1076 IleAlaGluLeuGlnAlaGlnIleAlaGluLeuLysAlaGlnLeuAlaLysLysGluGlu 1095

Qy       1490 GACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTA 1549

Db       1096 GluLeuGlnAlaAlaLeuAlaArgLeuGluAspGluThrSerGlnLysAsnAsnAlaLeu 1115

Qy       1550 CCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATG 1609
Db       1116 -----------------LysLysIleArgGluLeuGluSerHisIleSerAspLeuGln 1129

Qy       1610 TTCCCATCAGAATCCAAA---------CGAGAGGAAGATGAAGAATATTCTTGGGATTCT 1660
Db       1130 GluAspLeuGluSerGluLysAlaAlaArgAsnLysAlaGluLysGlnLysArgAspLeu 1149

Qy       1661 GGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTG------TGTATACCTGAGTCTATG 1714
Db       1150 SerGluGluLeuGlu---AlaLeuLysThrGluLeuGluAspThrLeuAspThrThrAla 1168

Qy       1715 TATCAGAAAGTAATGGAGATAAATAGAGAAGTAGAA-----------------GAGCTT 1756
Db       1169 ThrGlnGlnGluLeuArgAlaLysArgGluGlnGluValThrValLeuLysArgAlaLeu 1188

Qy       1757 CCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAA 1816
Db       1189 GluGluGluThrArgThrHisGluAlaGlnValGlnGluMetArgGlnLysHisThrGln 1208

Qy       1817 GCCTTTGAATTGAAGAATGAACAA------ACATTGAGAGCAGCTCAGATGTTCCCATCA 1870
Db       1209 AlaValGluGluLeuThrGluGlnLeuGluGlnPheLysArgAlaLysAlaAsnLeuAsp 1228

Qy       1871 GAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCCTGTGAGACG 1930
Db       1229 LysThrLysGlnThrLeuGluLysAspAsnAla---AspLeuAlaAsnGluIleArgSer 1247

Qy       1931 GTTTCACAG------AAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGAT 1984
Db       1248 LeuSerGlnAlaLysGlnAspVal--------------GluHisLysLysLys------ 1260

Qy       1985 ACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGT 2044
Db       1261 -----------LysLeuGlu---------------------------------------- 1263

Qy       2045 GGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAAA 2104
Db       1264 ---------ValGlnLeuGlnAspLeuGlnSerLysTyrSerAspGlyGluArgValArg 1280

Qy       2105 GCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTT 2164
Db       1281 ThrGluLeuAsnGluLysValHisLysLeuGlnIleGluValGluAsnValThrSerLeu 1300

Qy       2165 CCAAATAAAGCC------------TTAGAATTA---AAGGACAGAGAAACACTCAAAGCA 2209
Db       1301 LeuAsnGluAlaGluSerLysAsnIleLysLeuThrLysAspValAlaThrLeuGlySer 1320

Qy       2210 GAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCA 2269
Db       1321 GlnLeuGlnAspThrGlnGluLeuLeuGlnGluGluThrArgGlnLysLeuAsnValThr 1340

Qy       2270 AATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCA 2329
Db       1341 ThrLysLeuArgGlnLeuGluAspAspLysAsnSerLeuGlnGlnLeuAspGluGlu 1360

Qy       2330 ---GAATCCAAACAA-----------------------------------------AAGGAT 2347
Db       1361 ValGluAlaLysGlnAsnLeuGluArgHisIleSerThrLeuThrIleGlnLeuSerAsp 1380

Qy       2348 GATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGATGTG 2407
Db       1381 SerLysLysLysLeuGlnGluPheThrAlaThrValGluThrMetGluGluGly------ 1398

Qy       2408 TGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG 2467
Db       1399 ---------LysLysLysLeuGlnArgGluIleGluSerLeuThrGlnGlnPheGluGlu 1415

Qy       2468 ---TCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCA 2524
Db       1416 LysAlaAlaSerTyrAspLysLeuGluLysThrLysAsnArgLeuGlnGlnGluLeuAsp 1435
```

```
Qy       2525 AATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTA 2584
Db       1436 AspLeuValValAspLeuAspAsnGlnArgGlnLeu-----------ValSerAsnLeu 1451

Qy       2585 GAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAG---TCTACAAAA 2641
Db       1452 GluLysLysGlnLysLysPheAspGlnMetLeuAlaGluGluLysAsnIleSerSerLys 1471

Qy       2642 GTTGAGGAAGAC---------TTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTA 2692
Db       1472 TyrAlaAspGluArgAspArgAlaGluAlaGluAlaArgGluLysGluThrLysAlaLeu 1491

Qy       2693 ACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAAT 2752
Db       1492 SerLeuAlaArgAlaLeuGluGluAlaLeuGluAlaLysGluGluLeuGluArgThrAsn 1511

Qy       2753 AAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGAT------ 2806
Db       1512 LysMetLeuLysAlaGluMet---------GluAspLeuValSerSerLysAspAspVal 1528

Qy       2806 ------------------------------------------------------------ 2806
Db       1529 GlyLysAsnValHisGluLeuGluLysSerLysArgThrLeuGluGlnGlnValGluGlu 1548

Qy       2807 --------------------------TCTGAGATTATCTCTGTGAGTGATACACAG--- 2836
Db       1549 MetLysThrGlnLeuGluGluLeuGluAspGluLeuGlnAlaAlaGluAspAlaLysLeu 1568

Qy       2837 ------------AATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACA 2884
Db       1569 ArgLeuGluValAsnMetGlnAlaMet---LysSerGlnPheGluArgAspLeuGlnAla 1587

Qy       2885 ACAAATGGCAAAATAGAAGAGTCTCCTGAA---------AAGCCTTCTCACTTTGAGCCT 2935
Db       1588 ArgAspGluGlnAsnGluGluLysArgArgGlnLeuLeuLysGlnLeuHis---GluHis 1606

Qy       2936 GCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACA 2995
Db       1607 GluThrGluLeuGluAspGluArgLysGlnArgAlaLeuAlaAlaAlaAlaLysLysLys 1626

Qy       2996 TTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGA 3055
Db       1627 LeuGluValAspValLysAspLeuGluSerGlnValAsp---------SerAlaAsnLys 1643

Qy       3056 GGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAA 3115
Db       1644 AlaArgGlu---------GluAlaIleLysGlnLeuArgLysLeuGlnAlaGlnMetLys 1660

Qy       3116 AATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTA 3175
Db       1661 Asp------------TyrGlnArgAspLeuAspAspAlaArgAlaAlaArgGluGluIle 1676

Qy       3176 ---------------GAGAACCAAAAA------GCTAAATGGGAACAAGAGCTCTGCAGT 3214
Db       1677 PheAlaThrAlaArgGluAsnGluLysLysAlaLysAsnLeuGluAlaGluLeuIleGln 1696

Qy       3215 GTGAGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAA 3274
Db       1697 LeuGlnGluAspLeuAlaAlaAlaGluArgAlaArgLysGlnAlaAspLeuGluLysGlu 1716

Qy       3275 AAAATTAGACCCGAA--------------------------------GAGCAACTTAGG 3301
Db       1717 GluMetAlaGluGluLeuAlaSerAlaAsnSerGlyArgThrSerLeuGlnAspGluLys 1736

Qy       3302 AAAAAGTTAGAAGTGAAA---CACCAACTTGAACAGACTCTCAGAATACAAGATATAGAA 3358
Db       1737 ArgArgLeuGluAlaArgIleAlaGlnLeuGluGluGluLeuAspGluGluHisSerAsn 1756

Qy       3359 TTGAAAAGTGTAACAAGTAATTTGAATCAG---GTTTCTCACACTCATGAAAGTGAAAAT 3415
Db       1757 IleGluThrMetSerAspArgMetArgLysAlaValGlnGlnAlaGluGlnLeuAsnAsn 1776

Qy       3416 GATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAA---------------------- 3451
Db       1777 GluLeuAlaThrGluArgAlaThrAlaGlnLysAsnGluAsnAlaArgGlnGlnLeuGlu 1796

Qy       3451 ------------------------------------------------------------ 3451
Db       1797 ArgGlnAsnLysGluLeuArgSerLysLeuGlnGluMetGluGlyAlaValLysSerLys 1816

Qy       3452 ------------ATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAG 3499
Db       1817 PheLysSerThrIleAlaAlaLeuGluAlaLysIleAlaSerLeuGluGluGlnLeuGlu 1836

Qy       3500 GTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTT 3559
Db       1837 GlnGluAlaArgGluLysGlnAlaAlaAlaLysThrLeuArgGlnLysAspLysLysLeu 1856

Qy       3560 CAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAG 3619
Db       1857 LysAspAlaLeu---LeuGlnValGluAspGluArgLysGlnAlaGluGlnTyrLysAsp 1875

Qy       3620 CAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db       1876 GlnAlaGluLysGlyAsnLeuArgLeuLysGlnLeuLysArgGlnLeuGluGlu 1893
```

Search completed: July 15, 2004, 09:13:01
Job time : 167.051 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 09:09:52 ; Search time 53.6846 Seconds
(without alignments)
13162.486 Million cell updates/sec

Title: US-09-602-362E-26
Perfect score: 6448
Sequence: 1 caagagcttggcgatacaga.....tgacttctaaattgaaggaa 3673

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093626_22015/app_query.fasta_1.10325
-DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09602362_@CGN_1_1_248_@runat_15072004_093626_22015 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	347.5	5.4	1427	2	S22695	restin - human
2	343	5.3	1392	2	A43336	microtubule-vesicl
3	342	5.3	1790	2	S67593	transport protein
4	330	5.1	2663	1	S28261	centromere protein
5	324.5	5.0	1938	2	JC5421	smooth muscle myos
6	324.5	5.0	1972	2	JC5420	smooth muscle myos
7	319.5	5.0	2954	2	T14156	kinesin-related pr
8	318.5	4.9	2116	2	A26655	myosin heavy chain
9	315	4.9	1972	1	A41604	myosin heavy chain
10	315	4.9	3488	2	T34418	hypothetical prote
11	312.5	4.8	5327	2	T13564	microtubule-associ
12	311	4.8	3259	1	A56539	giantin - human
13	310	4.8	1961	1	A61231	myosin heavy chain
14	310	4.8	3225	2	I52300	giantin - human
15	302.5	4.7	3187	2	JC5837	364K Golgi complex
16	302	4.7	1004	2	A55142	myosin-light-chain
17	300.5	4.7	1313	2	F96673	hypothetical prote
18	300.5	4.7	2253	2	T30336	nuclear/mitotic ap
19	300	4.7	1979	1	S03166	myosin heavy chain
20	299.5	4.6	2677	2	A38194	desmoplakin I - hu
21	298	4.6	2057	2	S61477	myosin II heavy ch
22	297	4.6	2017	1	A36014	myosin heavy chain
23	294.5	4.6	1538	2	T29095	cardiac muscle fac
24	294	4.6	1939	2	T18372	repeat organellar
25	293.5	4.6	1690	2	T13030	microtubule bindin
26	293.5	4.6	1957	2	T38077	hypothetical coile
27	293	4.5	6642	2	T29757	protein UNC-89 - C
28	292.5	4.5	1410	1	A57013	early endosome ant
29	291.5	4.5	2331	2	T25410	hypothetical prote
30	290	4.5	2139	2	T18296	myosin heavy chain
31	289	4.5	1875	2	S38173	myosin-like protei
32	287.5	4.5	1738	2	T14867	interaptin - slime
33	286.5	4.4	1999	1	S21801	myosin heavy chain
34	286	4.4	993	2	S49461	synaptonemal compl
35	284.5	4.4	1558	2	B71603	RESA-H3 antigen PF
36	284.5	4.4	1727	2	T50073	myosin-like coiled
37	284	4.4	1642	2	T08880	NMDA receptor-bind
38	281	4.4	1390	2	S51364	sperm tail-specifi
39	280	4.3	1780	2	T17272	hypothetical prote
40	279	4.3	1295	2	T24587	hypothetical prote
41	279	4.3	1959	1	A33977	myosin heavy chain
42	277	4.3	1837	2	T41023	probable nuclear p
43	276.5	4.3	1356	2	S32763	kinectin 1 - human
44	276	4.3	1269	2	F84730	probable myosin he
45	275.5	4.3	1992	2	A47297	myosin heavy chain

ALIGNMENTS

RESULT 1
S22695
restin - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Nov-1999
C;Accession: S22695; S19853
R;Bilbe, G.; Delabie, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
EMBO J. 11, 2103-2113, 1992
A;Title: Restin: a novel intermediate filament-associated protein highly expressed in th
A;Reference number: S22695; MUID:92289675; PMID:1600942
A;Accession: S22695
A;Molecule type: mRNA
A;Residues: 1-1427 <BIL>
A;Cross-references: EMBL:X64838; NID:g35998; PIDN:CAA46050.1; PID:g35999
C;Keywords: cytoskeleton

Alignment Scores:

Pred. No.:	7.53e-12	Length:	1427
Score:	347.50	Matches:	304
Percent Similarity:	36.79%	Conservative:	207
Best Local Similarity:	21.89%	Mismatches:	510
Query Match:	5.39%	Indels:	368
DB:	2	Gaps:	67

US-09-602-362E-26 (1-3673) x S22695 (1-1427)

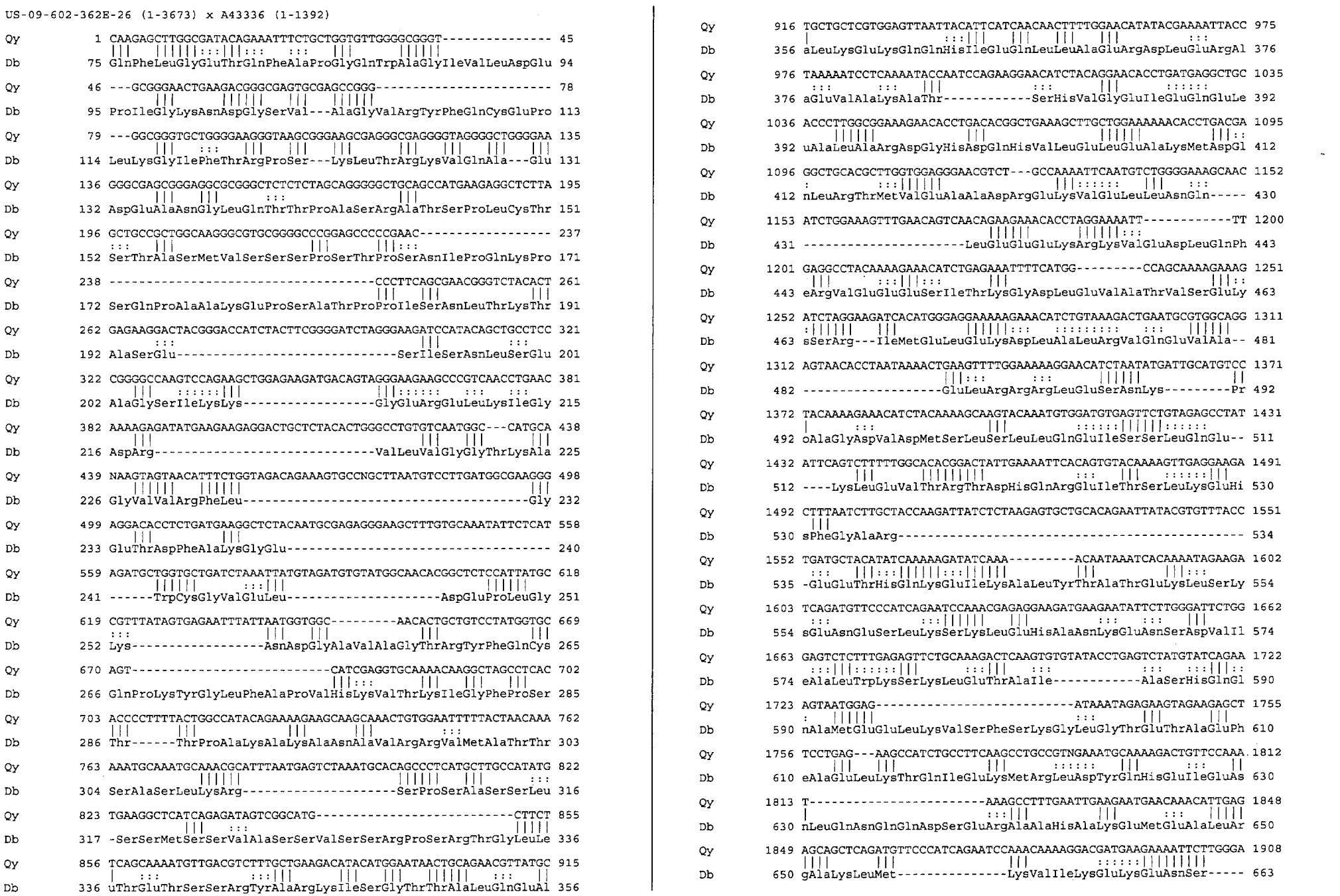
```
Qy          1 CAAGAGCTTGGCGATACAGAAATTTCTGCTGGTGTTGGGGCGGGT---------------- 45
              |||   ||||||:::|||:::   :::   |||       ||||||
Db         75 GlnPheLeuGlyGluThrGlnPheAlaProGlyGlnTrpAlaGlyIleValLeuAspGlu 94

Qy         46 ---GCGGGAACTGAAGACGGGCGAGTGCGAGCCGGG------------------------ 78
                 |||       ||||||   |||   ||||||
Db         95 ProIleGlyLysAsnAspGlySerVal---AlaGlyValArgTyrPheGlnCysGluPro 113

Qy         79 ---GGCGGGTGCTGGGGAAGGGTAAGCGGGAAGCGAGGGCGAGGGGTAGGGGCTGGGGAA 135
                 |||   :::   |||   |||   |||       |||   |||   |||   |||
Db        114 LeuLysGlyIlePheThrArgProSer---LysLeuThrArgLysValGlnAla---Glu 131
```

```
Db        739 aLysCysAsnGluGlnThrLysValIleAspAsnPheThrSerGlnLeuLysAlaThrGl 759

Qy       2095 AACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAA 2154

Db        759 u--------------------------GluLysLeuLeuAspLeuAspAlaLeuArgLy 770

Qy       2155 AGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAA-------- 2206

Db        770 sAlaSerSerGluGlyLysSerGluMetLysLysLeuArgGlnGlnLeuGluAlaAlaGl 790

Qy       2207 ----------------------------------GCAGAGTCTCCTGATAATGATGGTCT 2232

Db        790 uLysGlnIleLysHisLeuGluIleGluLysAsnAlaGluSerSerLysAlaSerSerIl 810

Qy       2233 TCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGA 2292

Db        810 eThrArgGluLeuGlnGlyArgGluLeuLysLeuThrAsnLeuGlnGluAsnLeuSerGl 830

Qy       2293 C------------AGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACA 2340

Db        830 uValSerGlnValLysGluThrLeuGluLysGluLeuGlnIleLeuLysGluLysPheAl 850

Qy       2341 AAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAA 2400

Db        850 aGluAlaSerGluGluAlaValSerValGlnArgSerMetGlnGluThrVal-------- 867

Qy       2401 TGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAA------TTCGATACCTTAAGTGG 2454

Db        868 ----------------AsnLysLeuHisGlnLysGluGluGlnPheAsnMetLeuSerSe 882

Qy       2455 AAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAAT 2514

Db        882 rAspLeuGluLys-----------------LeuArgGluAsnLeuAlaAspMetGluAl 896

Qy       2515 TTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGT 2574

Db        896 aLysPheArgGluLys------AspGluArgGluGluGlnLeuIleLysAlaLysGlu-- 913

Qy       2575 GAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAAT---TCACA 2631

Db        914 ----LysLeuGluAsnAspIleAlaGluIleMetLysMetSerGlyAspAsnSerSerGl 932

Qy       2632 GTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGT 2691

Db        932 nLeuThrLysMetAsnAspGluLeuArgLeu---------------------------- 942

Qy       2692 AACTGGACAACAGGAACGTGATATT---------------------------------- 2716

Db        943 ----------LysGluArgAspValGluGluLeuGlnLeuLysLeuThrLysAlaAsnGl 959

Qy       2717 -------GGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA-- 2767

Db        959 uAsnAlaSerPheLeuGlnLysSerIleGluAspMetThrValLysAlaGluGlnSerGl 979

Qy       2768 ----------------------GAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGA 2805

Db        979 nGlnGluAlaAlaLysLysHisGluGluGluLysLysGluLeuGluArgLysLeuSerAs 999

Qy       2806 TTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCT---GAGGCTAC 2862

Db        999 pLeuGlu------LysLysMetGluThrSerHisAsnGlnCysGlnGluLeuLysAlaAr 1017

Qy       2863 ATATCAAAAA---------GAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGA 2913

Db       1017 gTyrGluArgAlaThrSerGluThrLysThrLysHisGluGluIleLeuGlnAsnLeuGl 1037

Qy       2914 AAAG--------------CCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958

Db       1037 nLysThrLeuLeuAspThrGluAspLysLeuLysGlyAlaArgGlu---GluAsnSerGl 1056

Qy       2959 TCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCT 3018

Db       1056 yLeuLeuGlnGluLeuGluGluLeuArgLysGlnAlaAspLysAlaLysAlaAlaGlnTh 1076

Qy       3019 ATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAA 3078

Db       1076 rAlaGlu------AspAlaMetGlnIleMetGlu---------GlnMetThrLysGluLy 1091

Qy       3079 CTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAA 3138

Db       1091 sThrGluThrLeu---AlaSerLeuGluAspThrLysGlnThrAsnAlaLysLeuGlnAs 1110

Qy       3139 GGAACTGTCAGAAGCGAAAGAA------ATAAAATCACAGTTAGAGAACCAAAAAGCTAA 3192

Db       1110 nGluLeuAspThrLeuLysGluAsnAsnLeuLysAsnValGluGluLeuAsnLysSerLy 1130

Qy       3193 ATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAA---GAAGAGAAGAG 3249

Db       1130 s---------GluLeuLeuThrValGlu---------AsnGlnLysMetGluGluPheAr 1144

Qy       3250 AAGAAATGTCGATATATTAAAA---------------------------------------- 3271

Db       1144 gLysGluIleGluThrLeuLysGlnAlaAlaAlaGlnLysSerGlnGlnLeuSerAlaLe 1164

Qy       3272 -------GAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAG---------------- 3307

Db       1164 uGlnGluGluAsnValLysLeuAlaGluGluLeuGlyArgSerArgAspGluValThrSe 1184

Qy       3308 ---------------------------------------------TTAGAAGTGAAACA 3321

Db       1184 rHisGlnLysLeuGluGluGluArgSerValLeuAsnAsnGlnLeuLeuGluMetLysLy 1204

Qy       3322 CCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTT 3381

Db       1204 sArgGluSerLysPheIleLysAspAlaAspGluGluLysAlaSerLeuGlnLysSerIl 1224

Qy       3382 GAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTT 3441

Db       1224 eSer---IleThrSerAlaLeuLeuThrGluLysAspAlaGluLeuGluLys------Le 1241

Qy       3442 GAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGT 3501

Db       1241 uArgAsnGluValThrValLeuArgGlyGluAsnAlaSerAlaLysSerLeuHisSerVa 1261

Qy       3502 GAAGGAAAATAAATACTTTGAGGACATTAAGATT---TTACAAGAAAAGAATGCTGAACT 3558

Db       1261 lValGlnThrLeuGluSerAspLysValLysLeuGluLeuLysValLysAsnLeuGluLe 1281

Qy       3559 TCAAATGACCCTAAAACTGAAACAG 3583

Db       1281 uGlnLeuLysGluAsnLysArgGln 1289
```

RESULT 2
A43336
microtubule-vesicle linker CLIP-170 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C;Accession: A43336
R;Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
Cell 70, 887-900, 1992
A;Title: CLIP-170 links endocytic vesicles to microtubules.
A;Reference number: A43336; MUID:92405160; PMID:1356075
A;Accession: A43336
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1392 <PIE>
A;Cross-references: GB:M97501; NID:g180621; PIDN:AAA35693.1; PID:g180622

Alignment Scores:

Pred. No.:	1.33e-11	Length:	1392
Score:	343.00	Matches:	306
Percent Similarity:	35.95%	Conservative:	189
Best Local Similarity:	22.22%	Mismatches:	503
Query Match:	5.32%	Indels:	379
DB:	2	Gaps:	69


US-09-602-362E-26 (1-3673) x A43336 (1-1392)

```
Qy       1909 TTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCT----- 1963
Db        664 -------------LeuGluAlaIleArgSerLys---------LeuAspLysAlaGluAs 676

Qy       1964 -ACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGA 2022
Db        676 pGlnHisLeuValGluMetGluAspThrLeuAsnLysLeuGlnGluAlaGluIleLysVa 696

Qy       2023 T---------GGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT---CTTCCAAATAA 2070
Db        696 lLysGluLeuGluValLeuGlnAlaLysCysAsnGluGlnThrLysValIleAspAsnPh 716

Qy       2071 AGCCTTAGAATTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCT 2130
Db        716 eThrSerGlnLeuLysAlaThrGlu--------------------------GluLysLe 727

Qy       2131 TCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGA 2190
Db        727 uLeuAspLeuAspAlaLeuArgLysAlaSerSerGluGlyLysSerGluMetLysLysLe 747

Qy       2191 CAGAGAAACACTCAAA----------------------------------------GC 2208
Db        747 uArgGlnGlnLeuGluAlaAlaGluLysGlnIleLysHisLeuGluIleGluLysAsnAl 767

Qy       2209 AGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCC 2268
Db        767 aGluSerSerLysAlaSerSerIleThrArgGluLeuGlnGlyArgGluLeuLysLeuTh 787

Qy       2269 AAATAAAGCTTTAGAATTGAAGGAC------------AGAGAAACATTCAAAGCAGCTCA 2316
Db        787 rAsnLeuGlnGluAsnLeuSerGluValSerGlnValLysGluThrLeuGluLysGluLe 807

Qy       2317 GATGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAG 2376
Db        807 uGlnIleLeuLysGluLysPheAlaGluAlaSerGluGluAlaValSerValGlnArgSe 827

Qy       2377 TTTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGA 2436
Db        827 rMetGlnGluThrVal-----------------------AsnLysLeuHisGlnLysGl 839

Qy       2437 A------TTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCT 2490
Db        839 uGluGlnPheAsnMetLeuSerSerAspLeuGluLys-----------------LeuAr 853

Qy       2491 GAAGCCTACCTGTGGAATGAAAATTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAG 2550
Db        853 gGluAsnLeuAlaAspMetGluAlaLysPheArgGluLys------AspGluArgGluGl 871

Qy       2551 AGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAA 2610
Db        871 uGlnLeuIleLysAlaLysGlu------LysLeuGluAsnAspIleAlaGluIleMetLy 889

Qy       2611 ACCGACTACTGAAAAT---TCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTAC 2667
Db        889 sMetSerGlyAspAsnSerSerGlnLeuThrLysMetAsnAspGluLeuArgLeu----- 907

Qy       2668 CAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAACGTGATATT----------- 2716
Db        908 ---------------------------------LysGluArgAspValGluGluLeuGl 916

Qy       2717 ------------------------------GGCATTATTGAACGAGCTCCACAAGATCA 2745
Db        916 nLeuLysLeuThrLysAlaAsnGluAsnAlaSerPheLeuGlnLysSerIleGluAspMe 936

Qy       2746 AACAAATAAGATGCCCACATCA----------------------GAATTAGGAAGAAA 2781
Db        936 tThrValLysAlaGluGlnSerGlnGlnGluAlaAlaLysLysHisGluGluGluLysLy 956

Qy       2782 AGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTA 2841
Db        956 sGluLeuGluArgLysLeuSerAspLeuGlu------LysLysMetGluThrSerHisAs 974

Qy       2842 TGAGTGTTTACCT---GAGGCTACATATCAAAAA---------GAAATAAAGACAACAAA 2889
Db        974 nGlnCysGlnGluLeuLysAlaArgTyrGluArgAlaThrSerGluThrLysThrLysHi 994

Qy       2890 TGGCAAAATAGAAGAGTCTCCTGAAAAG--------------CCTTCTCACTTTGAGCC 2934
Db        994 sGluGluIleLeuGlnAsnLeuGlnLysThrLeuLeuAspThrGluAspLysLeuLysGl 1014

Qy       2935 TGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAAC 2994
Db        1014 yAlaArgGlu---GluAsnSerGlyLeuLeuGlnGluLeuGluGluLeuArgLysGlnAl 1033

Qy       2995 ATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAG 3054
Db        1033 aGluLysAlaLysAlaAlaGlnThrAlaGlu------AspAlaMetGlnIleMetGlu-- 1050

Qy       3055 AGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAA 3114
Db        1051 -------GlnMetThrLysGluLysThrGluThrLeu---AlaSerLeuGluAspThrLy 1067

Qy       3115 AAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAA------ATAAAATC 3168
Db        1067 sGlnThrAsnAlaLysLeuGlnAsnGluLeuAspThrLeuLysGluAsnAsnLeuLysAs 1087

Qy       3169 ACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTT 3228
Db        1087 nValGluGluLeuAsnLysSerLys---------GluLeuLeuThrValGlu-------- 1101

Qy       3229 AAATCAAGAA---GAAGAGAAGAGAAGAAATGTCGATATATTAAAA---------------- 3271
Db        1102 -AsnGlnLysMetGluGluPheArgLysGluIleGluThrLeuLysGlnAlaAlaAlaGl 1121

Qy       3272 ------------------------------GAAAAAATTAGACCCGAAGAGCAACTTAG 3300
Db        1121 nLysSerGlnGlnLeuSerAlaLeuGlnGluGluAsnValLysLeuAlaGluGluLeuGl 1141

Qy       3301 GAAAAAG---------------------------------------------------- 3307
Db        1141 yArgSerArgAspGluValThrSerHisGlnLysLeuGluGluGluArgSerValLeuAs 1161

Qy       3308 ----------TTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGATATAGA 3357
Db        1161 nAsnGlnLeuLeuGluMetLysLysArgGluSerLysPheIleLysAspAlaAspGluGl 1181

Qy       3358 ATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGA 3417
Db        1181 uLysAlaSerLeuGlnLysSerIleSer---IleThrSerAlaLeuLeuThrGluLysAs 1200

Qy       3418 TCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGC 3477
Db        1200 pAlaGluLeuGluLys------LeuArgAsnGluValThrValLeuArgGlyGluAsnAl 1218

Qy       3478 CACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATT-- 3535
Db        1218 aSerAlaLysSerLeuHisSerValValGlnThrLeuGluSerAspLysValLysLeuGl 1238

Qy       3536 -TTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAG 3583
Db        1238 uLeuLysValLysAsnLeuGluLeuGlnLeuLysGluAsnLysArgGln 1254
```

RESULT 3
S67593
transport protein USO1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D2552; protein YDL058w
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
C;Accession: S67593; A38455; S30782
R;Bloecker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67587
A;Accession: S67593
A;Molecule type: DNA
A;Residues: 1-1790 <BLO>
A;Cross-references: EMBL:Z74106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058w

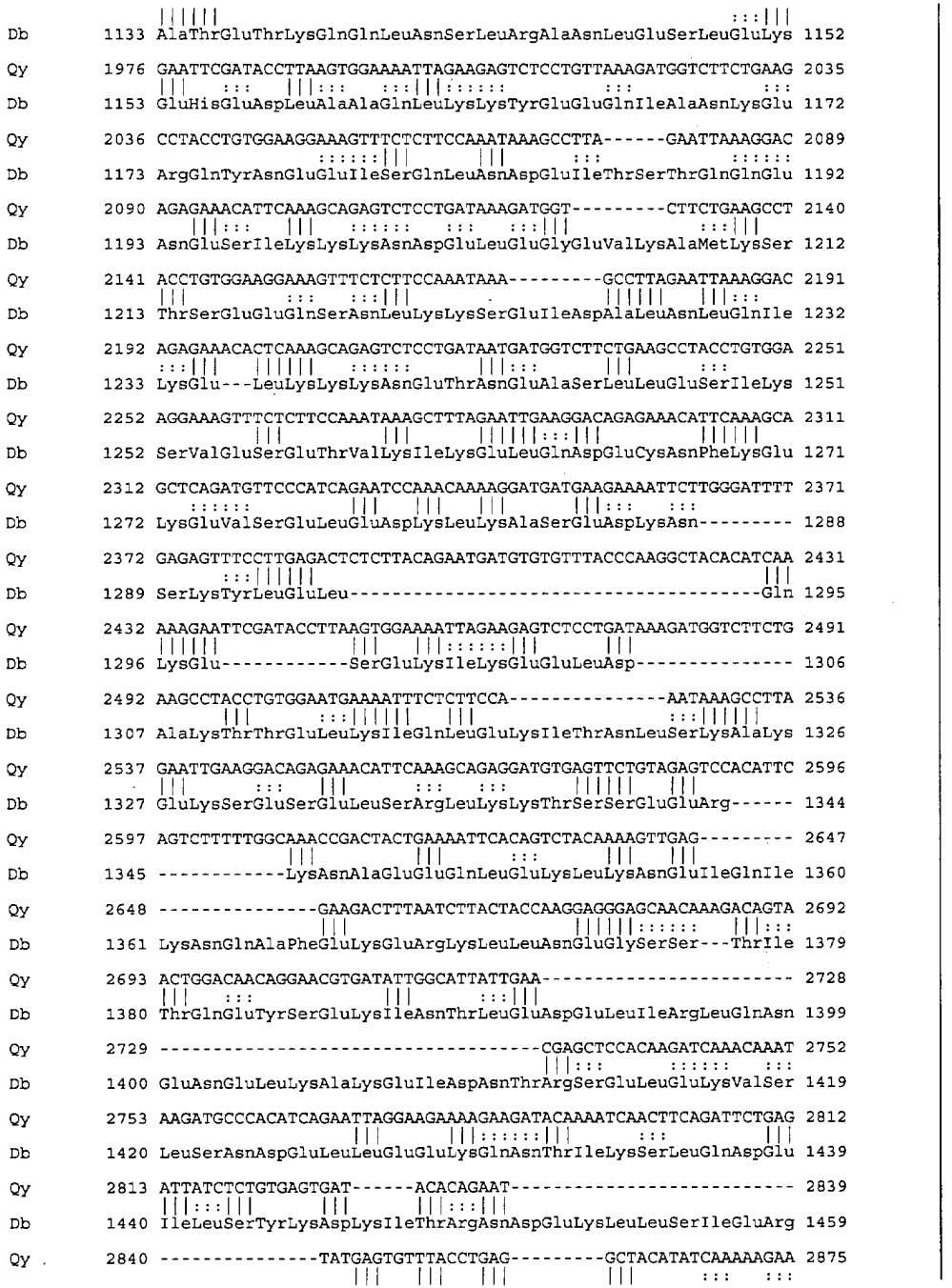
A;Experimental source: strain S288C
R;Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
J. Cell Biol. 113, 245-260, 1991
A;Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transp
A;Reference number: A38455; MUID:91185402; PMID:2010462
A;Accession: A38455
A;Molecule type: DNA
A;Residues: 1-389,'TA',392-724,'S',726-1790 <NAK>
A;Cross-references: GB:X54378; NID:g4777; PIDN:CAA38253.1; PID:g4778
A;Note: the authors translated the codon ACT for residue 768 as Ile
R;Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
submitted to the EMBL Data Library, February 1993
A;Description: An integrin analogue in Saccharomyces cerevisiae.
A;Reference number: S30782
A;Accession: S30782
A;Molecule type: DNA
A;Residues: 71-846,'E',848-923,'K',925-1252,'I',1254-1318,'V',1320-1460,'S',1462-1580,'S
A;Cross-references: EMBL:L03188
C;Genetics:
A;Gene: SGD:USO1; INT1
A;Cross-references: SGD:S0002216; MIPS:YDL058w
A;Map position: 4L
C;Keywords: coiled coil; transmembrane protein
F;326-342/Domain: transmembrane #status predicted <TM1>
F;394-410/Domain: transmembrane #status predicted <TM2>
F;617-633/Domain: transmembrane #status predicted <TM3>

Alignment Scores:

Pred. No.:	1.56e-11	Length:	1790
Score:	342.00	Matches:	261
Percent Similarity:	36.30%	Conservative:	180
Best Local Similarity:	21.48%	Mismatches:	422
Query Match:	5.30%	Indels:	352
DB:	2	Gaps:	55

US-09-602-362E-26 (1-3673) x S67593 (1-1790)

```
Qy        584 GTAGATGTGTATGGCAACACGGCTCTCCATTATGCCGTTTAT--------------AGT 628
              :::|||    |||        ::::::::::::   |||             :::
Db        686 LeuAspGluThrGlyLeuProLysValTyrPheSerThrTyrPheIleGlnLeuPheAsn 705

Qy        629 GAGAATTTATTAATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAAC 688
              ||||||:::      :::   |||   ||||||:::          |||      |||
Db        706 GluAsnIleTyrArgIleArgThrAlaLeuSerHisAspProAspGluGluProIleAsn 725

Qy        689 AAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAA 748
              |||   |||                                        :::   ||||||
Db        726 LysIleSerPhe-----------------------------------GluGluValGlu 733

Qy        749 TTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTC 808
                |||                                    :::   :::||||||   |||
Db        734 LysLeu---------------------------------GlnArgGlnCysThrLysLeu 742

Qy        809 ATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTT 868
                                                |||||          |||||
Db        743 --------------------LysGlyGluIleThrSer------LeuGlnThrGluThr 753

Qy        869 GACGTCTTTGCTGAAGACATA----------------------------CATGGAATA 898
              :::        |||::::::                            |||
Db        754 GluSerThrHisGluAsnLeuThrGluLysLeuIleAlaLeuThrAsnGluHisLysGlu 773

Qy        899 ACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAA 958
                  |||:::|||               :::|||      |||      |||   |||
Db        774 LeuAspGluLysTyrGlnIle---------LeuAsnSerSerHisSerSerLeuLysGlu 790

Qy        959 CATATACGAAAATTACCT---------AAAAATCCTCAAAATACCAATCCAGAAGGAACA 1009
              :::        |||           ||||||   :::::::::      |||   |||
Db        791 AsnPheSerIleLeuGluThrGluLeuLysAsnValArgAspSerLeuAspGluMetThr 810

Qy       1010 ------------------TCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGA 1051
                               |||        :::          |||   |||
Db        811 GlnLeuArgAspValLeuGluThrLysAspLysGluAsnGlnThrAlaLeuLeuGlu--- 829

Qy       1052 ACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTG 1111
                        :::|||   :::   |||      |||             |||
Db        830 -----------TyrLysSerThrIleHisLysGlnGluAspSerIleLysThrLeuGlu 845

Qy       1112 GAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGA------------ 1159
              :::|||           :::           ||||||     |||
Db        846 LysGlyLeuGluThrIleLeuSerGlnLysLysLysAlaGluAspGlyIleAsnLysMet 865

Qy       1160 ----------------------AAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAA 1195
                                     :::      |||:::   ||||||      :::
Db        866 GlyLysAspLeuPheAlaLeuSerArgGluMetGlnAlaValGluGluAsnCysLysAsn 885

Qy       1196 ATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCT 1255
              :::                                          |||||:::
Db        886 Leu-----------------------------------------GlnLysGluLysAsp 891

Qy       1256 AGGAAGATCACATGGGAGGAAAAAGAAACA---TCTGTAAAGACTGAATGCGTGGCAGGA 1312
              :::                 :::|||||||||   |||:::|||   :::      |||
Db        892 LysSerAsnValAsnHisGlnLysGluThrLysSerLeuLysGluAspIleAlaAlaLys 911

Qy       1313 GTAACACCTAATAAAACT---------------------------------------- 1330
              :::|||      |||
Db        912 IleThrGluIleLysAlaIleAsnGluAsnLeuGluGluMetLysIleGlnCysAsnAsn 931

Qy       1331 ----------------------------GAAGTTTTGGAA---------------AAA 1345
                                          |||::::::|||               :::
Db        932 LeuSerLysGluLysGluHisIleSerLysGluLeuValGluTyrLysSerArgPheGln 951

Qy       1346 GGAACATCTAATATGATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAAT 1405
                   |||::::::|||      |||::::::      :::   |||:::
Db        952 SerHisAspAsnLeuValAlaLysLeuThrGluLysLeuLysSerLeuAlaAsnAsnTyr 971

Qy       1406 GTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAAT 1465
                ||:::   :::                                        ||||||
Db        972 LysAspMetGlnAla-----------------------------------------GluAsn 978

Qy       1466 TCACAGTGTACAAAAGTTGAGGAAGAC--------------------------TTTAAT 1498
                        |||      |||:::                              |||
Db        979 GluSerLeuIleLysAlaValGluGluSerLysAsnGluSerSerIleGlnLeuSerAsn 998

Qy       1499 CTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCT 1558
              |||      ||||||   |||   |||      :::           :::
Db        999 LeuGlnAsnLysIleAspSerMetSerGlnGluLysGluAsnPheGlnIleGluArgGly 1018

Qy       1559 ACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGAT--------------- 1603
              :::  :::|||:::|||:::  :::        |||  |||
Db       1019 SerIleGluLysAsnIleGluGlnLeuLysLysThrIleSerAspLeuGluGlnThrLys 1038

Qy       1604 CAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGG 1663
              :::        |||:::|||        :::||||||      |||
Db       1039 GluGluIleIleSerLysSerAspSerSerLysAspGluTyrGluSer------GlnIle 1056

Qy       1664 AGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAA 1723
              ||||||   :::         :::|||                :::          |||
Db       1057 SerLeuLeuLysGluLysLeuGluThrAlaThrThrAlaAsnAspGluAsnValAsnLys 1076

Qy       1724 GTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCT 1783
              :::   |||:::   :::       |||||||||      :::   :::|||:::|||
Db       1077 IleSerGluLeuThrLysThrArgGluGluLeuGluAlaGluLeuAlaAlaTyrLys--- 1095

Qy       1784 GCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACA 1843
                                                  ||||||||||||
Db       1096 ------------------------------------AsnLeuLysAsnGlu------ 1100

Qy       1844 TTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCT 1903
                        ::::::    |||||     |||:::   :::||||||
Db       1101 ---LeuGluThrLysLeuGluThrSerGluLysAlaLeuLysGluValLysGlnAsnGlu 1119

Qy       1904 TGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA--- 1960
                                  |||   :::   ::::::   :::   |||   |||
Db       1120 --------------------GluHisLeuLysGluGluLysIleGlnLeuGluLysGlu 1132

Qy       1961 GCTACACAT-----------------------------------------------CAAAAA 1975
```



RESULT 4
S28261
centromere protein E - human
N;Alternate names: centromere 312K protein; kinesin-related protein CENP-E
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 19-Jan-2001
C;Accession: S28261
R;Yen, T.J.; Li, G.; Schaar, B.T.; Szilak, I.; Cleveland, D.W.
Nature 359, 536-539, 1992
A;Title: CENP-E is a putative kinetochore motor that accumulates just before mitosis.
A;Reference number: S28261; MUID:93024922; PMID:1406971

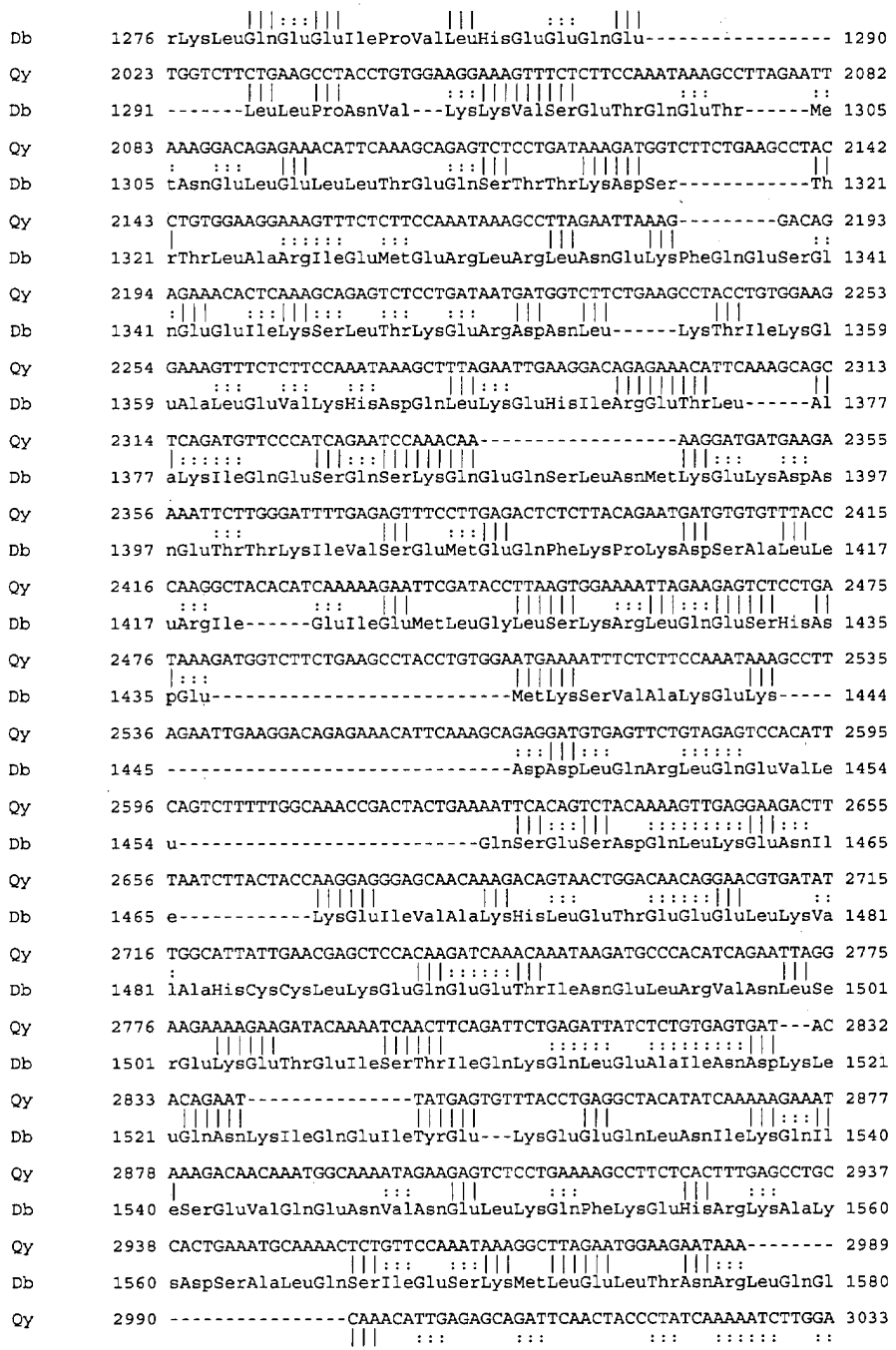
A;Accession: S28261
A;Molecule type: mRNA
A;Residues: 1-2663 <YEN>
A;Cross-references: EMBL:Z15005; NID:g29864; PIDN:CAA78727.1; PID:g29865
C;Genetics:
A;Gene: GDB:CENPE
A;Cross-references: GDB:361164; OMIM:117143
A;Map position: 4q24-4q25
C;Superfamily: centromere protein E; kinesin motor domain homology
C;Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop
F;7-335/Domain: kinesin motor domain homology <KMOT>
F;86-93/Region: nucleotide-binding motif A (P-loop)
F;486-2183/Domain: coiled coil #status predicted <COI>
F;92/Binding site: ATP (Lys) #status predicted

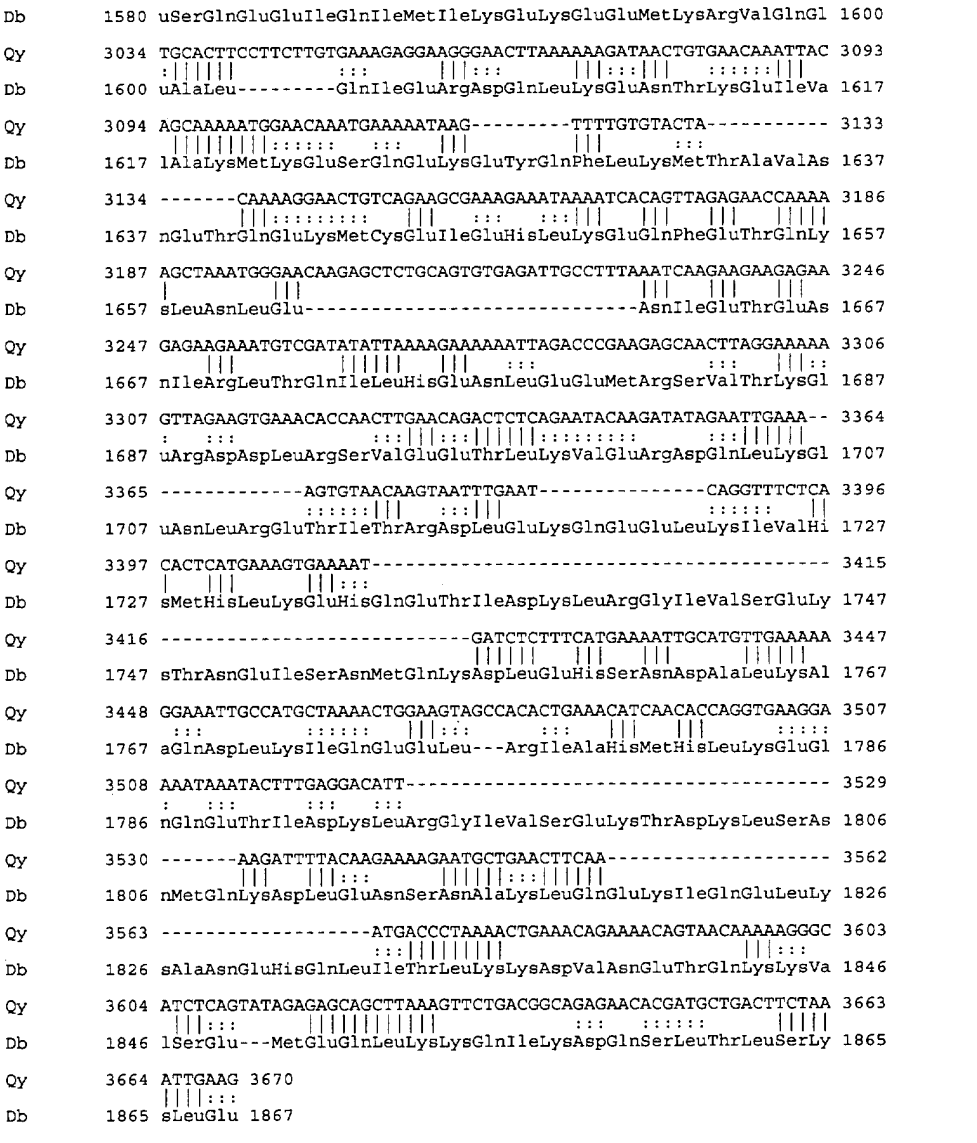
Alignment Scores:

Pred. No.:	7.56e-11	Length:	2663
Score:	330.00	Matches:	250
Percent Similarity:	37.53%	Conservative:	224
Best Local Similarity:	19.79%	Mismatches:	494
Query Match:	5.12%	Indels:	295
DB:	1	Gaps:	50

US-09-602-362E-26 (1-3673) x S28261 (1-2663)

```
Qy        478 AATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAA 537
Db        702 SerLeuIleAspGlyLysValProLysAspLeuLeuCysAsnLeuGluLeuGlu-GlyLy 721

Qy        538 GCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGG 597
Db        721 s---------------IleThrAspLeuGlnLysGluLeuAsn---LysGluValGluGl 735

Qy        598 CAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCT 657
Db        735 uAsnGluAlaLeuArg---------------GluGluValIleLeuLeuSerGluLeuLy 750

Qy        658 GTCCTATGGTGCAGTCATC---------------GAGGTGCAAAACAAGGCTAGCCTCAC 702
Db        750 sSerLeuProSerGluValGluArgLeuArgLysGluIleGlnAspLysSerGluGluLe 770

Qy        703 ACCCCTTTTACTGGCC--------------------------ATACAGAAAAGAAGCAA 735
Db        770 uHisIleIleThrSerGluLysAspLysLeuPheSerGluValValHisLysGluSerAr 790

Qy        736 G--------------------------CAAACTGTGGAATTTTTACTAACAAAAAATGC 768
Db        790 gValGlnGlyLeuLeuGluGluIleGlyLysThrLysAspAspLeuAlaThrThrGlnSe 810

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
Db        810 rAsnTyrLysSerThrAspGlnGluPheGlnAsnPheLysThrLeuHisMet------As 828

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
Db        828 pPheGluGlnLysTyrLysMetValLeuGluGluAsnGluArgMetAsnGlnGluIleVa 848

Qy        889 ACAT---GGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGA---------------GT 930
Db        848 lAsnLeuSerLysGluAlaGlnLysPheAspSerSerLeuGlyAlaLeuLysThrGluLe 868

Qy        931 TAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAA----------- 979
Db        868 uSerTyrLysThrGlnGluLeuGlnGluLysThrArgGluValGlnGluArgLeuAsnGl 888

Qy        980 ------------AATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGA 1026
Db        888 uMetGluGlnLeuLysGluGlnLeuGluAsnArgAspSerProLeuGlnThrValGluAr 908

Qy       1027 TGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTG------CTGGA 1080
Db        908 gGluLysThrLeuIleThrGluLysLeuGlnGlnThrLeuGluGluValLysThrLeuTh 928

Qy       1081 AAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCT 1140
Db        928 rGlnGluLysAspAspLeuLysGlnLeuGlnGluSerLeuGlnIleGluArgAspGlnLe 948

Qy       1141 GGGGAAAGCAACATCTGGAAAGTTTGAACAGTCA------ACAGAAGAAACACCTAGGAA 1194
Db        948 uLysSerAspIleHisAspThrValAsnMetAsnIleAspThrGlnGluGlnLeuArgAs 968

Qy       1195 AATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTT-----------TCATGGCCAGC 1242
Db        968 nAlaLeuGluSerLeuLysGlnHisGlnGluThrIleAsnThrLeuLysSerLysIleSe 988

Qy       1243 AAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATG 1302
Db        988 rGluGluValSerArgAsnLeuHisMetGluGluAsnThrGlyGluThrLysAspGluPh 1008

Qy       1303 C---------GTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATC 1353
Db       1008 eGlnGlnLysMetValGlyIle------AspLysLysGlnAspLeuGluAlaLysAsnTh 1026

Qy       1354 TAATATGATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGT 1413
Db       1026 rGlnThrLeuThrAlaAspValLysAsp--------------------AsnGluIleIl 1039

Qy       1414 GAGTTCTGTAGAGCCTATATTCAGTCTTTTGGCACACGG------------------- 1453
Db       1039 eGluGlnGlnArgLysIlePheSerLeuIleGlnGluLysAsnGluLeuGlnGlnMetLe 1059

Qy       1453 ------------------------------------------------------------ 1453
Db       1059 uGluSerValIleAlaGluLysGluGlnLeuLysThrAspLeuLysGluAsnIleGluMe 1079

Qy       1454 -ACTATTGAAAATTCACAGTGTACAAAA------------------------------ 1480
Db       1079 tThrIleGluAsnGlnGluGluLeuArgLeuLeuGlyAspGluLeuLysLysGlnGlnGl 1099

Qy       1481 ----GTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAA 1536
Db       1099 uIleValAlaGlnGluLysAsnHisAla---------IleLysLysGluGlyGluLeuSe 1116

Qy       1537 TTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAAT 1596
Db       1116 rArgThrCysAspArgLeuAlaGluValGluGluLysLeuLysGluLysSerGlnGlnLe 1136

Qy       1597 AGAAGATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGA 1656
Db       1136 uGlnGluLysGlnGlnGlnLeuLeuAsnValGlnGluGluMetSerGluMetGlnLysLy 1156

Qy       1657 TTCTGGGAGTCTC------TTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTC 1710
Db       1156 sIleAsnGluIleGluAsnLeuLysAsnGluLeuLysAsnLysGluLeuThrLeuGluHi 1176

Qy       1711 TATGTATCAGAAAGTAATGGAG-----------ATAAATAGAGAAGTAGAAGAGCTTCC 1758
Db       1176 sMetGluThrGluArgLeuGluLeuAlaGlnLysLeuAsnGluAsnTyrGluGluValLy 1196

Qy       1759 TGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGACTGTTCCAAAT----- 1813
Db       1196 sSerIleThrLysGluArgLysValLeuLysGluLeuGlnLysSerPheGluThrGluAr 1216

Qy       1814 ------------------------------AAAGCCTTTGAATTGAAGAATGAACAAAC 1842
Db       1216 gAspHisLeuArgGlyTyrIleArgGluIleGluAlaThrGlyLeuGlnThrLysGluGl 1236

Qy       1843 ATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTC 1902
Db       1236 uLeuLysIleAlaHisIleHisLeuLysGluHisGlnGluThrIleAspGluLeuArgAr 1256

Qy       1903 TTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGC 1962
Db       1256 gSerValSerGluLysThrAlaGlnIleIleAsnThrGlnAspLeuGluLysSerHisTh 1276

Qy       1963 TACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGA 2022
```





RESULT 5
JC5421
smooth muscle myosin heavy chain 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 02-Feb-2001
C;Accession: JC5421
R;Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A;Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A;Reference number: JC5420; MUID:97242182; PMID:9125171
A;Accession: JC5421

A;Molecule type: mRNA
A;Residues: 1-1938 <HAS>
A;Cross-references: DDBJ:D85924; NID:g1945079; PIDN:BAA19691.1; PID:g1945080
A;Experimental source: smooth muscle
C;Comment: This protein plays a role in smooth muscle cell contraction.
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: nucleotide binding; P-loop
F;88-771/Domain: myosin motor domain homology <MMOT>
F;178-185/Region: nucleotide-binding motif A (P-loop)

Alignment Scores:

Pred. No.:	1.47e-10	Length:	1938
Score:	324.50	Matches:	227
Percent Similarity:	37.07%	Conservative:	190
Best Local Similarity:	20.18%	Mismatches:	421
Query Match:	5.03%	Indels:	288
DB:	2	Gaps:	40

US-09-602-362E-26 (1-3673) x JC5421 (1-1938)

```
Qy        951 TTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACAT 1010
              :::|||   ::::::              :::|||::::::     |||:::|||
Db        834 TrpTrpArgLeuPheThrLysValLysProLeuGlnValThrArgGlnGluGluGlu 853

Qy       1011 CTACAG---------GAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACA 1061
              :::|||         |||  :::  ::::::          :::::::::      |||
Db        854 MetGlnAlaLysGluGluGluMetGlnLysIleThrGluArgGlnGlnLysAlaGluThr 873

Qy       1062 CGGCTGAAAGCTTGCTGGAAAAAACACCTG---------------------------- 1091
                ||||||          :::|||||||
Db        874 GluLeuLysGluLeuGluGlnLysHisThrGlnLeuAlaGluGluLysThrLeuGln 893

Qy       1092 ------------------ACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTC 1133
                               |||   |||:::|||         |||          :::
Db        894 GluGlnLeuGlnAlaGluThrGluLeuTyrAlaGluSerGluGluMetArgValArgLeu 913

Qy       1134 AATGTCTGGGGAAAGCAACATCTGGAA------------------------------- 1160
                     ||||||   ||||||
Db        914 AlaAla---LysLysGlnGluLeuGluGluIleLeuHisGluMetGluAlaArgLeuGlu 932

Qy       1161 ---------------------AGTTTGAACAGTCAACAGAAGAAACA-CCTAGGAAAATT 1198
                                  |||   :::::::::||||||      :::::::::
Db        933 GluGluGluAspArgArgGlnGlnLeuGlnAlaGluArgLysLysMetAlaGlnGlnMet 952

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
              |||      ::::::      |||:::   :::          |||   :::
Db        953 LeuAspLeuGluGluGlnLeuGluGluGluGluAlaAlaArgGlnLysLeuGlnLeuGlu 972

Qy       1259 AAGATCACATGGGAG---------------------------------------------- 1273
              |||:::|||   |||
Db        973 LysValThrAlaGluAlaLysIleLysLysLeuGluAspAspIleLeuValMetAspAsp 992

Qy       1274 --------------GAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
                           ||||||      :::      |||   |||:::   :::|||
Db        993 GlnAsnSerLysLeuSerLysGluArgLysLeuLeuGluGluArgValSerAspLeuThr 1012

Qy       1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
                 |||      |||      ||||||          |||:::          :::|||
Db       1013 ThrAsnLeuAlaGluGluGluGluLys---AlaLysAsnLeuThrLysLeuLysSerLys 1031

Qy       1379 GAAACATCTACAAAAGCAAGTACAAATGTG---------------------------- 1408
                    |||      :::      |||
Db       1032 HisGluSerMetIleSerGluLeuGluValArgLeuLysLysGluGluLysSerArgGln 1051

Qy       1409 ---------------------------------GATGTGAGTTCTGTA---GAGCCTATA 1432
                                              |||   |||      |||   |||
Db       1052 GluLeuGluLysLeuLysArgLysLeuGluGlyAspAlaSerAspPheHisGluGlnIle 1071

Qy       1433 TTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGAC 1492
                 |||                        |||      |||   ||||||:::
Db       1072 AlaAspLeuGlnAlaGlnIleAlaGluLeuLysMetGlnLeuAlaLysLysGluGluGlu 1091

Qy       1493 TTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCT 1552
                                    |||      :::   :::      ||||||          |||
Db       1092 LeuGlnAlaAlaLeuAlaArgLeuAspGluGluIleAlaGlnLysAsnAsnAlaLeu--- 1110

Qy       1553 GATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTC 1612
                              |||   |||:::   :::          |||   |||
Db       1111 --------------LysLysIleArgGluLeuGluGlyHisIleSerAspLeuGlnGlu 1125

Qy       1613 CCATCAGAATCCAAA---------CGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGG 1663
                  :::|||:::          |||   :::   |||:::          |||   |||
Db       1126 AspLeuAspSerGluArgAlaAlaArgAsnLysAlaGluLysGlnLysArgAspLeuGly 1145

Qy       1664 AGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTG------TGTATACCTGAGTCTATGTAT 1717
                     |||   :::   |||||:::::          :::      :::
Db       1146 GluGluLeuGlu---AlaLeuLysThrGluLeuGluAspThrLeuAspSerThrAlaThr 1164

Qy       1718 CAGAAAGTAATGGAGATAAATAGAGAA------------------------------- 1744
              |||:::   :::          ||||||
Db       1165 GlnGlnGluLeuArgAlaLysArgGluGlnGluValThrValLeuLysLysAlaLeuAsp 1184

Qy       1744 ------------------------------------------------------------ 1744

Db       1185 GluGluThrArgSerHisGluAlaGlnValGlnGluMetArgGlnLysHisThrGlnAla 1204

Qy       1745 GTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTN------------GAA 1792
              ||||||||||||   |||:::          ||||||   |||          :::
Db       1205 ValGluGluLeuThrGluGlnLeuGluGlnPheLysArgAlaLysAlaAsnLeuAspLys 1224

Qy       1793 ATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCA 1852
                 ::::::|||:::      :::      :::|||      |||      :::
Db       1225 SerLysGlnThrLeuGluLysGluAsnAlaAspLeuAlaGlyGluLeuArgValLeuGly 1244

Qy       1853 GCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCT 1912
                              |||   |||:::|||      |||          |||
Db       1245 GlnAlaLysGlnGluValGluHisLysLysLysLysLeuGluValGlnLeuGlnAspLeu 1264

Qy       1913 GAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACAT--- 1969
              :::|||   |||          :::          |||          |||
Db       1265 GlnSerLysCysSerAspGlyGluArgAlaArgAlaGluLeuSerAspLysValHisLys 1284

Qy       1970 ---CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGT 2026
                 |||   |||   ::::::::::::|||   |||   |||:::
Db       1285 LeuGlnAsnGluValGluSerValThrGlyMetLeuAsnGluAla--------------- 1299

Qy       2027 CTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTA--- 2083
                                                        ||||||::::::|||
Db       1300 ----------------------------------GluGlyLysAlaIleLysLeuAla 1307

Qy       2084 AAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2143
              ||||||      :::      ::::::      |||          ||||||:::
Db       1308 LysAspValAlaSerLeuGlySerGlnLeuGlnAspThrGlnGluLeuLeuGlnGluGlu 1327

Qy       2144 TGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTC 2203
                 :::|||:::::::::      |||      :::||||:::|||
Db       1328 ThrArgGlnLysLeuAsnValSerThrLysLeuArgGlnLeuGluAspGluArgAsnSer 1347

Qy       2204 AAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT 2263
                 :::      |||   :::      |||          |||   ||||||
Db       1348 LeuGlnAspGlnLeuAspGluGluMetGluAlaLysGlnAsnLeuGluArgHisValSer 1367

Qy       2264 CTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTC 2323
                 |||      |||
Db       1368 ThrLeuAsnIleGlnLeu--------------------------------------- 1373

Qy       2324 CCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTT 2383
              |||:::||||||:::|||                    ||||||   |||   :::
Db       1374 ---SerAspSerLysLysLys--------------LeuGlnAspPheAlaSerThrIle 1387

Qy       2384 GAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGAT 2443
              |||   :::  :::                |||          |||||||||   :::
Db       1388 GluValMetGluGluGly--------------LysLysArgLeuGlnLysGluMetGlu 1402

Qy       2444 ACCTTAAGTGGAAAATTAGAAGAG----------------------TCTCCTGATAAA 2479
                 ||||||   :::   ||||||                              |||
```

```
Db       1403 GlyLeuSerGlnGlnTyrGluGluLysAlaAlaAlaTyrAspLysLeuGluLysThrLys 1422

Qy       2480 GATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCC------ 2533

Db       1423 AsnArgLeuGlnGlnGluLeuAspAspLeuValValAspLeuAspAsnGlnArgGlnLeu 1442

Qy       2534 ---------TTAGAATTGAAGGACAGA-------------------------------- 2551

Db       1443 ValSerAsnLeuGluLysLysGlnLysLysPheAspGlnLeuLeuAlaGluGluLysAsn 1462

Qy       2552 --------------------GAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCC 2590

Db       1463 IleSerSerLysTyrAlaAspGluArgAspArgAlaGluAlaGluAlaArgGluLysGlu 1482

Qy       2591 ACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAA 2650

Db       1483 ThrLysAlaLeuSerLeuAlaArgAlaLeuGlu------GluAlaLeuGluAlaLysGlu 1500

Qy       2651 GACTTTAATCTTACTACCAAGGAG-----------------------GGAGCAACAAAG 2686

Db       1501 GluLeuGluArgThrAsnLysMetLeuLysAlaGluMetGluAspLeuValSerSerLys 1520

Qy       2687 ACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAA 2746

Db       1521 AspAspValGlyLysAsnValHisGluLeuGluLysSerLysArgAlaLeuGluThrGln 1540

Qy       2747 ACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGAT 2806

Db       1541 MetGluGluMetLysThrGlnLeuGluGluSerGluAspAspValGlnAlaThrGluAsp 1560

Qy       2807 TCTGAGATT---ATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACA 2863

Db       1561 AlaLysLeuArgLeuGluVal-----------AsnMetGlnAlaLeu---LysGlyGln 1575

Qy       2864 TATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCT---------GAA 2914

Db       1576 PheGluArgAspLeuGlnAlaArgAspGluGlnAsnGluGluLysArgArgGlnLeuGln 1595

Qy       2915 AAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTA 2974

Db       1596 ArgGlnLeuHis---GluTyrGluThrGluLeuGluAspGluArgLysGlnArgAlaLeu 1614

Qy       2975 GAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGAT 3034

Db       1615 AlaAlaAlaAlaLysLysLysLeuGluGlyAspLeuLysAspLeuGluLeuGlnAlaAsp 1634

Qy       3035 GCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACA 3094

Db       1635 ---------SerAlaIleLysGlyArgGlu---------GluAlaIleLysGlnLeuArg 1648

Qy       3095 GCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCG 3154

Db       1649 LysLeuGlnAlaGlnMetLysAsp-----------PheGlnArgGluLeuAspAspAla 1664

Qy       3155 AAAGAAATAAAATCACAGTTA--------------GAGAACCAAAAA------GCTAAA 3193

Db       1665 ArgAlaSerArgAspGluIlePheAlaThrSerLysGluAsnGluLysLysAlaLysSer 1684

Qy       3194 TGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGA 3253

Db       1685 LeuGluAlaAspLeuMetGlnLeuGlnGluAspLeuAlaAlaAlaGluArgAlaArgLys 1704

Qy       3254 AATGTCGATATATTAAAAGAAAAAATTAGACCCGAA---------------------- 3289

Db       1705 GlnAlaAspLeuGluLysGluGluLeuAlaGluGluLeuAlaSerSerLeuSerGlyArg 1724

Qy       3290 ---------GAGCAACTTAGGAAAAAGTTAGAAGTGAAA---CACCAACTTGAACAGACT 3337

Db       1725 AsnThrLeuGlnAspGluLysArgArgLeuGluAlaArgIleAlaGlnLeuGluGluGlu 1744

Qy       3338 CTCAGAATACAAGATATAGAATTGAAAAGTGTA--------------------ACAAGT 3376

Db       1745 LeuGluGluGluGlnGlyAsnMetGluAlaMetSerAspArgValArgLysAlaThrLeu 1764

Qy       3377 AATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAAT--- 3433

Db       1765 GlnAlaGluGlnLeuSerAsnGluLeuAlaThrGluArgSerThrAlaGlnLysAsnGlu 1784

Qy       3433 ------------------------------------------------------------ 3433

Db       1785 SerAlaArgGlnGlnLeuGluArgGlnAsnLysGluLeuArgSerLysLeuGlnGluVal 1804

Qy       3434 --------------TGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCC 3478

Db       1805 GluGlyAlaValLysAlaLysLeuLysSerThrValAlaAlaLeuGluAlaLysIleAla 1824

Qy       3479 ACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTA 3538

Db       1825 GlnLeuGluGluGlnValGluGlnGluAlaArgGluLysGlnAlaAlaThrLysSerLeu 1844

Qy       3539 CAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAA 3598

Db       1845 LysGlnLysAspLysLysLeuLysGluValLeu---LeuGlnValGluAspGluArgLys 1863

Qy       3599 AGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACT 3658

Db       1864 MetAlaGluGlnTyrLysGluGlnAlaGluLysGlyAsnThrLysValLysGlnLeuLys 1883

Qy       3659 TCTAAATTGAAGGAA 3673

Db       1884 ArgGlnLeuGluGlu 1888
```

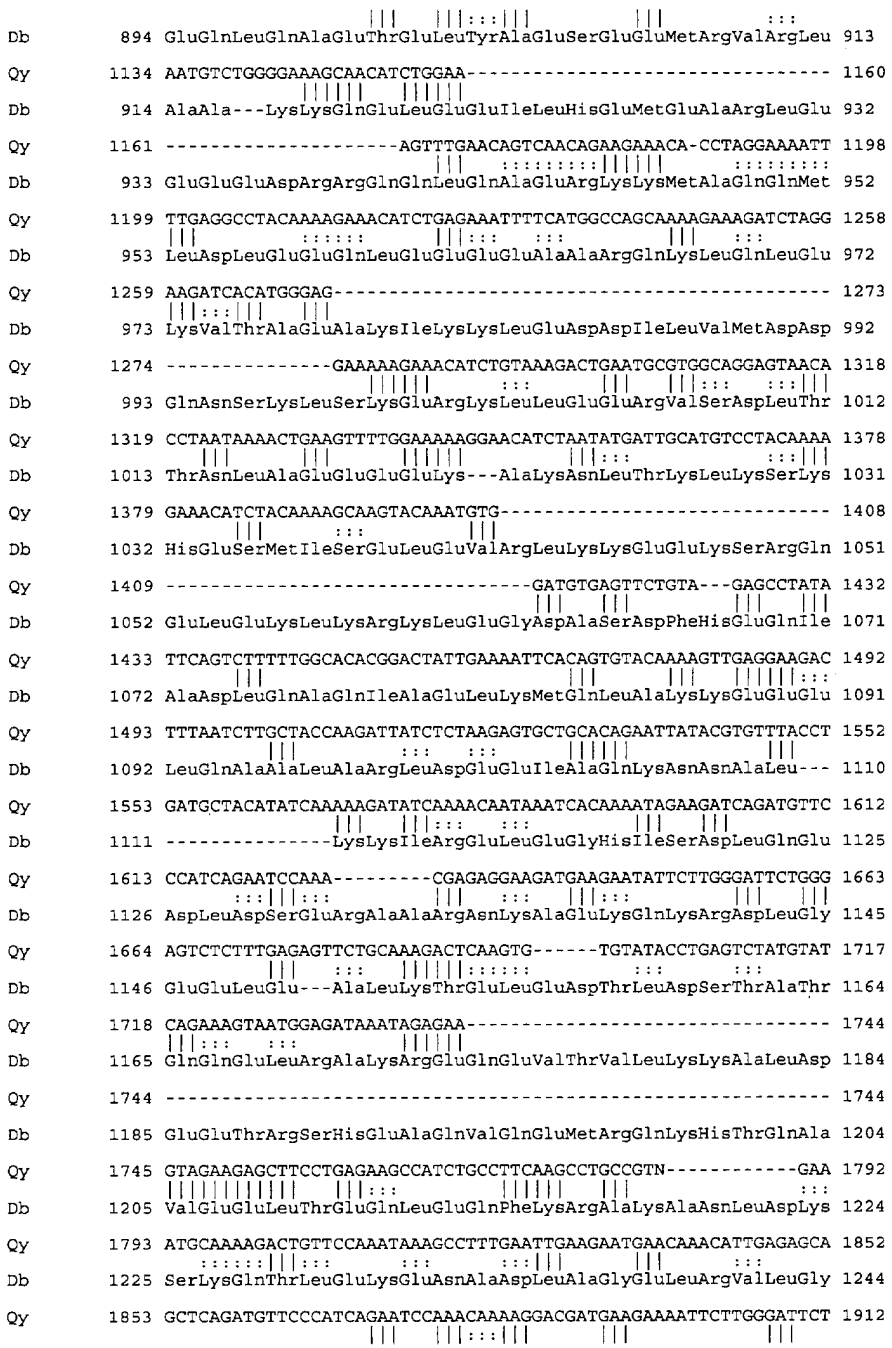
RESULT 6
JC5420
smooth muscle myosin heavy chain 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 02-Feb-2001
C;Accession: JC5420
R;Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A;Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A;Reference number: JC5420; MUID:97242182; PMID:9125171
A;Accession: JC5420
A;Molecule type: mRNA
A;Residues: 1-1972 <HAS>
A;Cross-references: DDBJ:D85923; NID:g1945077; PIDN:BAA19690.1; PID:g1945078
A;Experimental source: smooth muscle
C;Comment: This protein plays a role in smooth muscle cell contraction.
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: nucleotide binding; P-loop
F;88-771/Domain: myosin motor domain homology <MMOT>
F;178-185/Region: nucleotide-binding motif A (P-loop)

Alignment Scores:

Pred. No.:	1.48e-10	Length:	1972
Score:	324.50	Matches:	227
Percent Similarity:	37.07%	Conservative:	190
Best Local Similarity:	20.18%	Mismatches:	421
Query Match:	5.03%	Indels:	288
DB:	2	Gaps:	40

US-09-602-362E-26 (1-3673) x JC5420 (1-1972)

```
Qy        951 TTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACAT 1010

Db        834 TrpTrpArgLeuPheThrLysValLysProLeuLeuGlnValThrArgGlnGluGluGlu 853

Qy       1011 CTACAG---------GAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACA 1061

Db        854 MetGlnAlaLysGluGluGluMetGlnLysIleThrGluArgGlnGlnLysAlaGluThr 873

Qy       1062 CGGCTGAAAGCTTGCTGGAAAAAACACCTG---------------------------- 1091

Db        874 GluLeuLysGluLeuGluGlnLysHisThrGlnLeuAlaGluGluLysThrLeuLeuGln 893

Qy       1092 -----------------ACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTC 1133
```


Db 894 GluGlnLeuGlnAlaGluThrGluLeuTyrAlaGluSerGluGluMetArgValArgLeu 913
Qy 1134 AATGTCTGGGGAAAGCAACATCTGGAA---------- 1160
Db 914 AlaAla---LysLysGlnGluLeuGluGluIleLeuHisGluMetGluAlaArgLeuGlu 932
Qy 1161 ----------AGTTTGAACAGTCAACAGAAGAAACA-CCTAGGAAAATT 1198
Db 933 GluGluGluAspArgArgGlnGlnLeuGlnAlaGluArgLysLysMetAlaGlnGlnMet 952
Qy 1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
Db 953 LeuAspLeuGluGluGlnLeuGluGluGluGluAlaAlaArgGlnLysLeuGlnLeuGlu 972
Qy 1259 AAGATCACATGGGAG---------- 1273
Db 973 LysValThrAlaGluAlaLysIleLysLysLeuGluAspAspIleLeuValMetAspAsp 992
Qy 1274 ----------GAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACA 1318
Db 993 GlnAsnSerLysLeuSerLysGluArgLysLeuLeuGluGluArgValSerAspLeuThr 1012
Qy 1319 CCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAA 1378
Db 1013 ThrAsnLeuAlaGluGluGluGluLys---AlaLysAsnLeuThrLysLeuLysSerLys 1031
Qy 1379 GAAACATCTACAAAAGCAAGTACAAATGTG---------- 1408
Db 1032 HisGluSerMetIleSerGluLeuGluValArgLeuLysLysGluGluLysSerArgGln 1051
Qy 1409 ----------GATGTGAGTTCTGTA---GAGCCTATA 1432
Db 1052 GluLeuGluLysLeuLysArgLysLeuGluGlyAspAlaSerAspPheHisGluGlnIle 1071
Qy 1433 TTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGAC 1492
Db 1072 AlaAspLeuGlnAlaGlnIleAlaGluLeuLysMetGlnLeuAlaLysLysGluGluGlu 1091
Qy 1493 TTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCT 1552
Db 1092 LeuGlnAlaAlaLeuAlaArgLeuAspGluGluIleAlaGlnLysAsnAsnAlaLeu--- 1110
Qy 1553 GATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTC 1612
Db 1111 ----------LysLysIleArgGluLeuGluGlyHisIleSerAspLeuGlnGlu 1125
Qy 1613 CCATCAGAATCCAAA---------CGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGG 1663
Db 1126 AspLeuAspSerGluArgAlaAlaArgAsnLysAlaGluLysGlnLysArgAspLeuGly 1145
Qy 1664 AGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTG------TGTATACCTGAGTCTATGTAT 1717
Db 1146 GluGluLeuGlu---AlaLeuLysThrGluLeuGluAspThrLeuAspSerThrAlaThr 1164
Qy 1718 CAGAAAGTAATGGAGATAAATAGAGAA---------- 1744
Db 1165 GlnGlnGluLeuArgAlaLysArgGluGlnGluValThrValLeuLysLysAlaLeuAsp 1184
Qy 1744 ---------- 1744
Db 1185 GluGluThrArgSerHisGluAlaGlnValGlnGluMetArgGlnLysHisThrGlnAla 1204
Qy 1745 GTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTN----------GAA 1792
Db 1205 ValGluGluLeuThrGluGlnLeuGluGlnPheLysArgAlaLysAlaAsnLeuAspLys 1224
Qy 1793 ATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCA 1852
Db 1225 SerLysGlnThrLeuGluLysGluAsnAlaAspLeuAlaGlyGluLeuArgValLeuGly 1244
Qy 1853 GCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCT 1912
Db 1245 GlnAlaLysGlnGluValGluHisLysLysLysLeuGluValGlnLeuGlnAspLeu 1264
Qy 1913 GAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACAT--- 1969
Db 1265 GlnSerLysCysSerAspGlyGluArgAlaArgAlaGluLeuSerAspLysValHisLys 1284
Qy 1970 ---CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGT 2026
Db 1285 LeuGlnAsnGluValGluSerValThrGlyMetLeuAsnGluAla---------- 1299
Qy 2027 CTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTA--- 2083
Db 1300 ----------GluGlyLysAlaIleLysLeuAla 1307
Qy 2084 AAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2143
Db 1308 LysAspValAlaSerLeuGlySerGlnLeuGlnAspThrGlnGluLeuLeuGlnGluGlu 1327
Qy 2144 TGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTC 2203
Db 1328 ThrArgGlnLysLeuAsnValSerThrLysLeuArgGlnLeuGluAspGluArgAsnSer 1347
Qy 2204 AAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCT 2263
Db 1348 LeuGlnAspGlnLeuAspGluGluMetGluAlaLysGlnAsnLeuGluArgHisValSer 1367
Qy 2264 CTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTC 2323
Db 1368 ThrLeuAsnIleGlnLeu---------- 1373
Qy 2324 CCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTT 2383
Db 1374 ---SerAspSerLysLysLys----------LeuGlnAspPheAlaSerThrIle 1387
Qy 2384 GAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGAT 2443
Db 1388 GluValMetGluGluGly----------LysLysArgLeuGlnLysGluMetGlu 1402
Qy 2444 ACCTTAAGTGGAAAATTAGAAGAG----------TCTCCTGATAAA 2479
Db 1403 GlyLeuSerGlnGlnTyrGluGluLysAlaAlaAlaTyrAspLysLeuGluLysThrLys 1422
Qy 2480 GATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCC------ 2533
Db 1423 AsnArgLeuGlnGlnGluLeuAspAspLeuValValAspLeuAspAsnGlnArgGlnLeu 1442
Qy 2534 ---------TTAGAATTGAAGGACAGA---------- 2551
Db 1443 ValSerAsnLeuGluLysLysGlnLysLysPheAspGlnLeuLeuAlaGluGluLysAsn 1462
Qy 2552 ----------GAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCC 2590
Db 1463 IleSerSerLysTyrAlaAspGluArgAspArgAlaGluAlaGluAlaArgGluLysGlu 1482
Qy 2591 ACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTTGAGGAA 2650
Db 1483 ThrLysAlaLeuSerLeuAlaArgAlaLeuGlu------GluAlaLeuGluAlaLysGlu 1500
Qy 2651 GACTTTAATCTTACTACCAAGGAG----------GGAGCAACAAAG 2686
Db 1501 GluLeuGluArgThrAsnLysMetLeuLysAlaGluMetGluAspLeuValSerSerLys 1520
Qy 2687 ACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAA 2746
Db 1521 AspAspValGlyLysAsnValHisGluLeuGluLysSerLysArgAlaLeuGluThrGln 1540
Qy 2747 ACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGAT 2806
Db 1541 MetGluGluMetLysThrGlnLeuGluGluSerGluAspAspValGlnAlaThrGluAsp 1560
Qy 2807 TCTGAGATT---ATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACA 2863
Db 1561 AlaLysLeuArgLeuGluVal----------AsnMetGlnAlaLeu---LysGlyGln 1575

```
Qy       2864 TATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCT---------GAA 2914
Db       1576 PheGluArgAspLeuGlnAlaArgAspGluGlnAsnGluGluLysArgArgGlnLeuGln 1595

Qy       2915 AAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTA 2974
Db       1596 ArgGlnLeuHis---GluTyrGluThrGluLeuGluAspGluArgLysGlnArgAlaLeu 1614

Qy       2975 GAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGAT 3034
Db       1615 AlaAlaAlaAlaLysLysLysLeuGluGlyAspLeuLysAspLeuGluLeuGlnAlaAsp 1634

Qy       3035 GCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACA 3094
Db       1635 ---------SerAlaIleLysGlyArgGlu---------GluAlaIleLysGlnLeuArg 1648

Qy       3095 GCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCG 3154
Db       1649 LysLeuGlnAlaGlnMetLysAsp-----------PheGlnArgGluLeuAspAspAla 1664

Qy       3155 AAAGAAATAAAATCACAGTTA--------------GAGAACCAAAAA------GCTAAA 3193
Db       1665 ArgAlaSerArgAspGluIlePheAlaThrSerLysGluAsnGluLysLysAlaLysSer 1684

Qy       3194 TGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGA 3253
Db       1685 LeuGluAlaAspLeuMetGlnLeuGlnGluAspLeuAlaAlaAlaGluArgAlaArgLys 1704

Qy       3254 AATGTCGATATATTAAAAGAAAAAATTAGACCCGAA---------------------- 3289
Db       1705 GlnAlaAspLeuGluLysGluGluLeuAlaGluGluLeuAlaSerSerLeuSerGlyArg 1724

Qy       3290 ---------GAGCAACTTAGGAAAAAGTTAGAAGTGAAA---CACCAACTTGAACAGACT 3337
Db       1725 AsnThrLeuGlnAspGluLysArgArgLeuGluAlaArgIleAlaGlnLeuGluGluGlu 1744

Qy       3338 CTCAGAATACAAGATATAGAATTGAAAAGTGTA--------------------ACAAGT 3376
Db       1745 LeuGluGluGluGlnGlyAsnMetGluAlaMetSerAspArgValArgLysAlaThrLeu 1764

Qy       3377 AATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAAT--- 3433
Db       1765 GlnAlaGluGlnLeuSerAsnGluLeuAlaThrGluArgSerThrAlaGlnLysAsnGlu 1784

Qy       3433 ------------------------------------------------------------ 3433
Db       1785 SerAlaArgGlnGlnLeuGluArgGlnAsnLysGluLeuArgSerLysLeuGlnGluVal 1804

Qy       3434 ---------------TGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCC 3478
Db       1805 GluGlyAlaValLysAlaLysLeuLysSerThrValAlaAlaLeuGluAlaLysIleAla 1824

Qy       3479 ACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTA 3538
Db       1825 GlnLeuGluGluGlnValGluGlnGluAlaArgGluLysGlnAlaAlaThrLysSerLeu 1844

Qy       3539 CAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAA 3598
Db       1845 LysGlnLysAspLysLysLeuLysGluValLeu---LeuGlnValGluAspGluArgLys 1863

Qy       3599 AGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACT 3658
Db       1864 MetAlaGluGlnTyrLysGluGlnAlaGluLysGlyAsnThrLysValLysGlnLeuLys 1883

Qy       3659 TCTAAATTGAAGGAA 3673
Db       1884 ArgGlnLeuGluGlu 1888
```

RESULT 7
T14156
kinesin-related protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: T14156
R;Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A;Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromoso
A;Reference number: Z17893; MUID:98028574; PMID:9363944
A;Accession: T14156
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2954 <WOO>
A;Cross-references: EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAC60300.1
C;Genetics:
A;Gene: XCENP-E
C;Superfamily: centromere protein E; kinesin motor domain homology

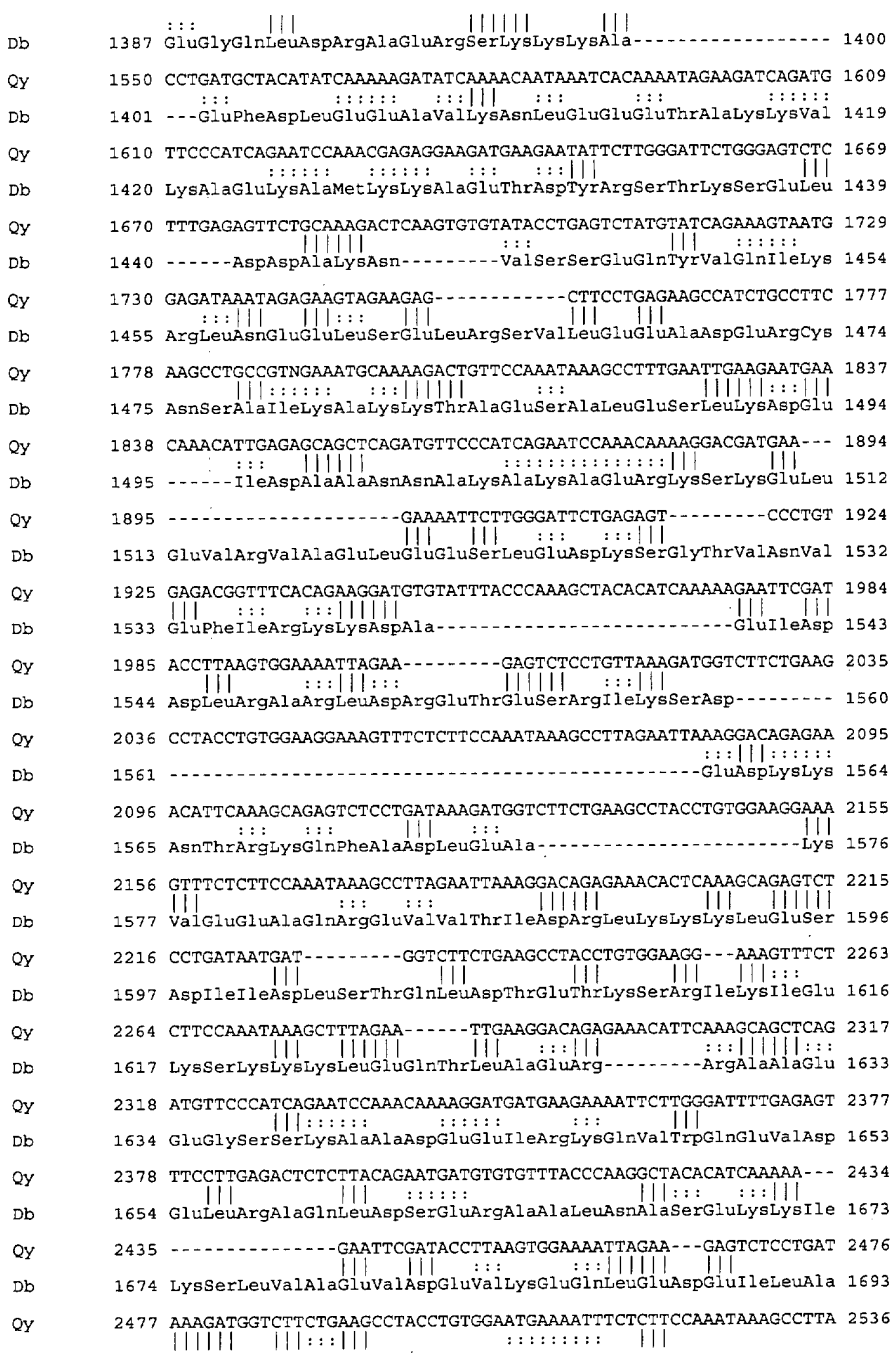
Alignment Scores:

Pred. No.:	2.93e-10	Length:	2954
Score:	319.50	Matches:	251
Percent Similarity:	37.32%	Conservative:	211
Best Local Similarity:	20.27%	Mismatches:	397
Query Match:	4.96%	Indels:	380
DB:	2	Gaps:	54

US-09-602-362E-26 (1-3673) x T14156 (1-2954)

```
Qy        759 CAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAG---------------- 803
Db       1168 GlnValIleAspLeuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLysSer 1187

Qy        804 ------------CCCTCATGC-----------------------------------TTG 815
Db       1188 AspLeuGlnLysProLysGlnAspLeuGluGluGlyGluValLysLeuLeuLeuGluMet 1207

Qy        816 CCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATG--------- 866
Db       1208 GluLeuLeuLysGlyHisLeuThrAspSerGlnLeuSerIleGluLysLeuGlnLeuGlu 1227

Qy        867 ---------TTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCTG 917
Db       1228 AsnLeuGluValThrGluLysLeuGlnThrLeuGlnGluGluMetLysAsnIleThrIle 1247

Qy        918 CTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTA 977
Db       1248 GluArgAsnGluLeuGlnThr---------AsnPheGluAspLeuLysAlaGluHisAsp 1264

Qy        978 AAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCAC 1037
Db       1265 SerLeu------------------LysGlnAspLeuSerGluAsnIleGluGlnSerIle 1278

Qy       1038 CCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAA------AAACACCTG 1091
Db       1279 GluThrGlnAspGlu------------LeuArgAlaAlaGlnGluGluLeuArgGluGln 1294

Qy       1092 ACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGG---GGAAAG 1148
Db       1295 LysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIleSerSer 1314

Qy       1149 CAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTA 1208
Db       1315 ProAsnHisAspAlaValAlaAsnGlnGluLysValSerLeuGlyGluValAsnSerLeu 1334

Qy       1209 CAA---------------------------------------------------------- 1211
Db       1335 GlnSerGluMetLeuArgGlyGluArgAspGluLeuGlnThrSerCysLysAlaLeuVal 1354

Qy       1212 ------------------AAGAAACATCTGAGAAATTTTCATGGCCAGCAAAA-GAAAGA 1252
Db       1355 SerGluLeuGluLeuLeuArgAlaHisValLysSerValGluGlyGluAsnLeuGluIle 1374

Qy       1253 TCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGA 1312
Db       1375 ThrLysLysLeuAsnGlyLeuGluLysGluIleLeuGlyLysSerGlu------------ 1390
```

```
Db       2047 GluAlaGluLysGluIleSerGluAlaThrAsnGluIleLysAsnLeuThrAlaLysIle 2066

Qy       2921 TCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGG 2980

Db       2067 SerSerLeuGluGluGluIleLeuGlnAsnAlaSerIleLeuAsnGluAlaVal------ 2084

Qy       2981 AAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTT 3040

Db       2085 SerGluArgGluAsnLeuArgHisSerLysGlnGlnLeuValSerGluLeuGluGlnLeu 2104

Qy       3041 CCTTCTTGTGAAAGAGGAAGG--------------GAACTTAAAAAAGATAACTGTGAA 3085

Db       2105 SerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArgGluLysAspGlu 2124

Qy       3086 CAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTG 3145

Db       2125 AlaValAsn---LysIleAlaSerLeuAlaGluGluIleLysIleLeuThrLysGluMet 2143

Qy       3146 TCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAG 3205

Db       2144 AspGluPheArgAspSerLysGluSerLeuGlnGluGlnSerSerHisLeuSerGluGlu 2163

Qy       3206 CTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTC---GAT 3262

Db       2164 LeuCysThrTyrLysThrGluLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnAsn 2183

Qy       3263 ATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAGTTA---GAAGTGAAA 3319

Db       2184 LysLeuAlaGluLysValLysGluValAspGluLeuLeuGlnHisLeuSerSerLeuLys 2203

Qy       3320 CACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAAT 3379

Db       2204 GluGlnLeuAspGlnIle------------GlnMetGluLeuArg--------------- 2214

Qy       3380 TTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATG 3439

Db       2215 ---AsnGluLysLeuArgAsnTyrGlu------------LeuCysGluLysMetAspIle 2229

Qy       3440 TTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAG 3499

Db       2230 MetGluLysGluIleSerValLeuArgLeu------------MetGlnAsnGluProGln 2245

Qy       3500 GTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTT 3559

Db       2246 GlnGluGluAspAspValAlaGluArgMetAspIleLeuGluSerArgAsnGlnGluIle 2265

Qy       3560 CAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAG 3619

Db       2266 --------------------GlnGluLeuMetGluLysIleSerAlaValTyrSerGlu 2278

Qy       3620 CAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673

Db       2279 GlnHisThrLeuLeuSerSer------------LeuSerSerGluLeuGlnLys 2292

RESULT 8
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Dictyostelium discoideum
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 19-Apr-2002
C;Accession: A26655; A24728; S00250
R;Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A;Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium disc
A;Reference number: A26655; MUID:87092266; PMID:3540939
A;Accession: A26655
A;Molecule type: DNA
A;Residues: 1-2116 <WAR>
A;Cross-references: GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1; PID:g167835
R;DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A;Reference number: A24728; MUID:86016788; PMID:3901008
A;Accession: A24728
A;Molecule type: mRNA
A;Residues: 2035-2116 <DEL>
R;Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A;Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A;Reference number: S00250; MUID:88112226; PMID:2828113
A;Accession: S00250
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1734-1893 <WAG>
C;Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phos
F;1-818/Domain: globular head <HED>
F;89-747/Domain: myosin motor domain homology <MMOT>
F;179-186/Region: nucleotide-binding motif A (P-loop)
F;819-2116/Domain: alpha-helical rod <ROD>

Alignment Scores:
Pred. No.:                3.2e-10        Length:          2116
Score:                    318.50         Matches:         210
Percent Similarity:       37.96%         Conservative:    181
Best Local Similarity:    20.39%         Mismatches:      376
Query Match:              4.94%          Indels:          263
DB:                       2              Gaps:            44

US-09-602-362E-26 (1-3673) x A26655 (1-2116)

Qy        857 CAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCT 916

Db       1201 LysLysLysValGluLeuAspLeuGluAspLysSerAlaGlnLeuAlaGluGluThrAla 1220

Qy        917 GCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTA--- 973

Db       1221 AlaLysGlnAlaLeuAspLysLeuLysLysLysLeuGluGlnGluLeuSerGluValGln 1240

Qy        974 ------------CCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCT------ACA 1015

Db       1241 ThrGlnLeuSerGluAlaAsnAsnLysAsnValAsnSerAspSerThrAsnLysHisLeu 1260

Qy       1016 GGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTG 1075

Db       1261 GluThrSerPheAsnAsnLeuLysLeuGluLeuGluAlaGluGlnLysAlaLysGlnAla 1280

Qy       1076 CTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAA 1135

Db       1281 LeuGluLys---------------LysArgLeu------GlyLeuGluSerGluLeuLys 1293

Qy       1136 TGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAA 1195

Db       1294 HisValAsn---------------------GluGlnLeuGluGluGlu------------ 1302

Qy       1196 ATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCT 1255

Db       1303 ------LysLysGlnLysGluSerAsnGluLys-------------------------- 1311

Qy       1256 AGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTA 1315

Db       1312 ArgLysValAspLeuGluLysGluValSerGluLeuLysAspGlnIleGluGluGluVal 1331

Qy       1316 ACACCTAATAAA------ACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGT 1369

Db       1332 AlaSerLysLysAlaValThrGluAlaLysAsnLysLysGluSerGluLeuAspGluIle 1351

Qy       1370 CCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCT 1429

Db       1352 LysArgGlnTyrAlaAspValValSerSerArgAsp---------LysSerValGluGln 1368

Qy       1430 ATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAA 1489

Db       1369 LeuLysThrLeuGlnAlaLysAsn------GluGluLeuArgAsnThrAlaGluGluAla 1386

Qy       1490 GACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTA 1549
```


Db 1387 GluGlyGlnLeuAspArgAlaGluArgSerLysLysLysAla---------- 1400
Qy 1550 CCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATG 1609
Db 1401 ---GluPheAspLeuGluGluAlaValLysAsnLeuGluGluGluThrAlaLysLysVal 1419
Qy 1610 TTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTC 1669
Db 1420 LysAlaGluLysAlaMetLysLysAlaGluThrAspTyrArgSerThrLysSerGluLeu 1439
Qy 1670 TTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATG 1729
Db 1440 ------AspAspAlaLysAsn---------ValSerSerGluGlnTyrValGlnIleLys 1454
Qy 1730 GAGATAAATAGAGAAGTAGAAGAG------------CTTCCTGAGAAGCCATCTGCCTTC 1777
Db 1455 ArgLeuAsnGluGluLeuSerGluLeuArgSerValLeuGluGluAlaAspGluArgCys 1474
Qy 1778 AAGCCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAA 1837
Db 1475 AsnSerAlaIleLysAlaLysLysThrAlaGluSerAlaLeuGluSerLeuLysAspGlu 1494
Qy 1838 CAAACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAA--- 1894
Db 1495 ------IleAspAlaAlaAsnAsnAlaLysAlaLysAlaGluArgLysSerLysGluLeu 1512
Qy 1895 ----------GAAAATTCTTGGGATTCTGAGAGT---------CCCTGT 1924
Db 1513 GluValArgValAlaGluLeuGluGluSerLeuGluAspLysSerGlyThrValAsnVal 1532
Qy 1925 GAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGAT 1984
Db 1533 GluPheIleArgLysLysAspAla----------GluIleAsp 1543
Qy 1985 ACCTTAAGTGGAAAATTAGAA---------GAGTCTCCTGTTAAAGATGGTCTTCTGAAG 2035
Db 1544 AspLeuArgAlaArgLeuAspArgGluThrGluSerArgIleLysSerAsp--------- 1560
Qy 2036 CCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAA 2095
Db 1561 ----------GluAspLysLys 1564
Qy 2096 ACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAA 2155
Db 1565 AsnThrArgLysGlnPheAlaAspLeuGluAla----------Lys 1576
Qy 2156 GTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCT 2215
Db 1577 ValGluGluAlaGlnArgGluValValThrIleAspArgLeuLysLysLysLeuGluSer 1596
Qy 2216 CCTGATAATGAT---------GGTCTTCTGAAGCCTACCTGTGGAAGG---AAAGTTTCT 2263
Db 1597 AspIleIleAspLeuSerThrGlnLeuAspThrGluThrLysSerArgIleLysIleGlu 1616
Qy 2264 CTTCCAAATAAAGCTTTAGAA------TTGAAGGACAGAGAAACATTCAAAGCAGCTCAG 2317
Db 1617 LysSerLysLysLysLeuGluGlnThrLeuAlaGluArg---------ArgAlaAlaGlu 1633
Qy 2318 ATGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGT 2377
Db 1634 GluGlySerSerLysAlaAlaAspGluGluIleArgLysGlnValTrpGlnGluValAsp 1653
Qy 2378 TTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAA--- 2434
Db 1654 GluLeuArgAlaGlnLeuAspSerGluArgAlaAlaLeuAsnAlaSerGluLysLysIle 1673
Qy 2435 --------------GAATTCGATACCTTAAGTGGAAAATTAGAA---GAGTCTCCTGAT 2476
Db 1674 LysSerLeuValAlaGluValAspGluValLysGluGlnLeuGluAspGluIleLeuAla 1693
Qy 2477 AAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTA 2536
Db 1694 LysAspLysLeuValLysAlaLysArgAlaLeuGluValGluLeuGluGluValArgAsp 1713
Qy 2537 GAATTGAAGGACAGAGAAACATTCAAAGCAGAG---------- 2569
Db 1714 GlnLeuGluGluGluGluGluAspSerArgSerGluLeuGluAspSerLysArgArgLeuThr 1733
Qy 2570 ---GATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAAT 2626
Db 1734 ThrGluValGluAspIleLysLysLysTyrAsp----------AlaGluVal 1747
Qy 2627 TCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAG 2686
Db 1748 GluGlnAsnThrLysLeuAspGlu----------AlaLysLys 1758
Qy 2687 ACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAA 2746
Db 1759 LysLeuThrAsp----------AspValAspThrLeuLysLysGlnLeuGluAspGlu 1774
Qy 2747 ACAAATAAGATGCCCACATCAGAATTAGGAAGAAAA----------GAAGAT 2788
Db 1775 LysLysLysLeuAsnGluSerGluArgAlaLysLysArgLeuGluSerGluAsnGluAsp 1794
Qy 2789 ACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGT 2848
Db 1795 PheLeuAlaLysLeuAspAlaGluValLysAsnArgSerArgAlaGluLys--------- 1811
Qy 2849 TTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCT 2908
Db 1812 ---AspArgLysLysTyrGluLysAspLeuLysAspThrLysTyrLysLeuAsnAsp--- 1829
Qy 2909 CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAA 2968
Db 1830 ----------GluAlaAlaThrLysThrGlnThrGluIleGlyAlaAla 1842
Qy 2969 GGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATC 3028
Db 1843 LysLeuGlu---------- 1845
Qy 3029 TTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAA 3088
Db 1846 ----------AspGlnIleAspGlu 1850
Qy 3089 ATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCA 3148
Db 1851 LeuArgSerLysLeuGluGlnGluAlaLysAlaThrGlnAlaAspLys--------- 1867
Qy 3149 GAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTC 3208
Db 1868 SerLysLysThrLeuGluGlyGluIleAspAsnLeuArgAlaGlnIleGluAspGlu--- 1886
Qy 3209 TGCAGTGTGAGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGA---------AATGTC 3259
Db 1887 ------GlyLysIleLysMetArgLeuGluLysGluLysArgAlaLeuGluGlyGluLeu 1904
Qy 3260 GATATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAA 3319
Db 1905 GluGluLeuArgGluThrValGluGluAlaGlu----------AspSerLys 1918
Qy 3320 CACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAAT 3379
Db 1919 SerGluAlaGluGlnSerLysArgLeuValGluLeuGluLeuGluAspAlaArgArgAsn 1938
Qy 3380 TTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATG 3439
Db 1939 LeuGlnLysGluIleAspAlaLysGluIleAlaGluAsp---------AlaLysSerAsn 1955
Qy 3440 TTGAAAAAGGAAATTGCC---------- 3457
Db 1956 LeuGlnArgGluIleValGluAlaLysGlyArgLeuGluGluGluSerIleAlaArgThr 1975
Qy 3458 ----------ATGCTAAAACTGGAAGTAGCCACACTGAAACATCAA--- 3493
Db 1976 AsnSerAspArgSerArgLysArgLeuGluAlaGluIleAspAlaLeuThrAlaGlnVal 1995

```
Qy       3494 ------------------------CACCAGGTGAAGGAAAATAAATACTTTGAG---GAC 3526
                                      |||:::|||||||||||      |||   :::
Db       1996 AspAlaGluGlnLysAlaLysAsnGlnGlnIleLysGluAsnLysLysIleGluThrGlu 2015

Qy       3527 ATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAA 3586
              :::|||     ::::::|||      |||   :::   |||        :::
Db       2016 LeuLysGluTyrArgLysLysPheGlyGluSerGluLysThrLysThrLysGluPheLeu 2035

Qy       3587 ACAGTAACAAAAAGGGCATCTCAGTATAGA-----------------------GAGCAG 3622
                 |||   |||     :::   |||:::                        ||||||
Db       2036 ValValGluLysLeuGluThrAspTyrLysArgAlaLysLysGluAlaAlaAspGluGln 2055

Qy       3623 CTTAAAGTTCTGACGGCAGAGAACACGATG 3652
                 :::   ||||||   ||||||   :::
Db       2056 GlnGlnArgLeuThrValGluAsnAspLeu 2065
```

RESULT 9
A41604
myosin heavy chain, smooth muscle, long splice form - rabbit
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C;Accession: A41604; A33501
R;Babij, P.; Kelly, C.; Periasamy, M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10676-10680, 1991
A;Title: Characterization of a mammalian smooth muscle myosin heavy-chain gene: complete
A;Reference number: A41604; MUID:92073350; PMID:1961735
A;Accession: A41604
A;Molecule type: mRNA
A;Residues: 1-1972 <BAB>
A;Cross-references: GB:M77812
R;Nagai, R.; Kuro-o, M.; Babij, P.; Periasamy, M.
J. Biol. Chem. 264, 9734-9737, 1989
A;Title: Identification of two types of smooth muscle myosin heavy chain isoforms by cDN
A;Reference number: A33501; MUID:89255535; PMID:2722872
A;Accession: A33501
A;Molecule type: mRNA
A;Residues: 1455-1972 <NAG>
A;Cross-references: GB:J04833; NID:g165519; PIDN:AAA31407.1; PID:g165520
A;Experimental source: smooth muscle
A;Note: examination by Southern blotting for the regions of difference between this isof
ative splicing
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated
F;88-771/Domain: myosin motor domain homology <MMOT>
F;178-185/Region: nucleotide-binding motif A (P-loop)
F;559-572/Region: actin binding #status predicted
F;633-647/Region: actin binding #status predicted
F;844-1938/Domain: coiled coil #status predicted <COI>
F;844-1284/Region: S2
F;1285-1972/Region: light meromyosin
F;1939-1972/Domain: carboxyl-terminal <CBT>
F;129/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F;184/Binding site: ATP (Lys) #status predicted
F;701,711/Active site: Cys #status predicted

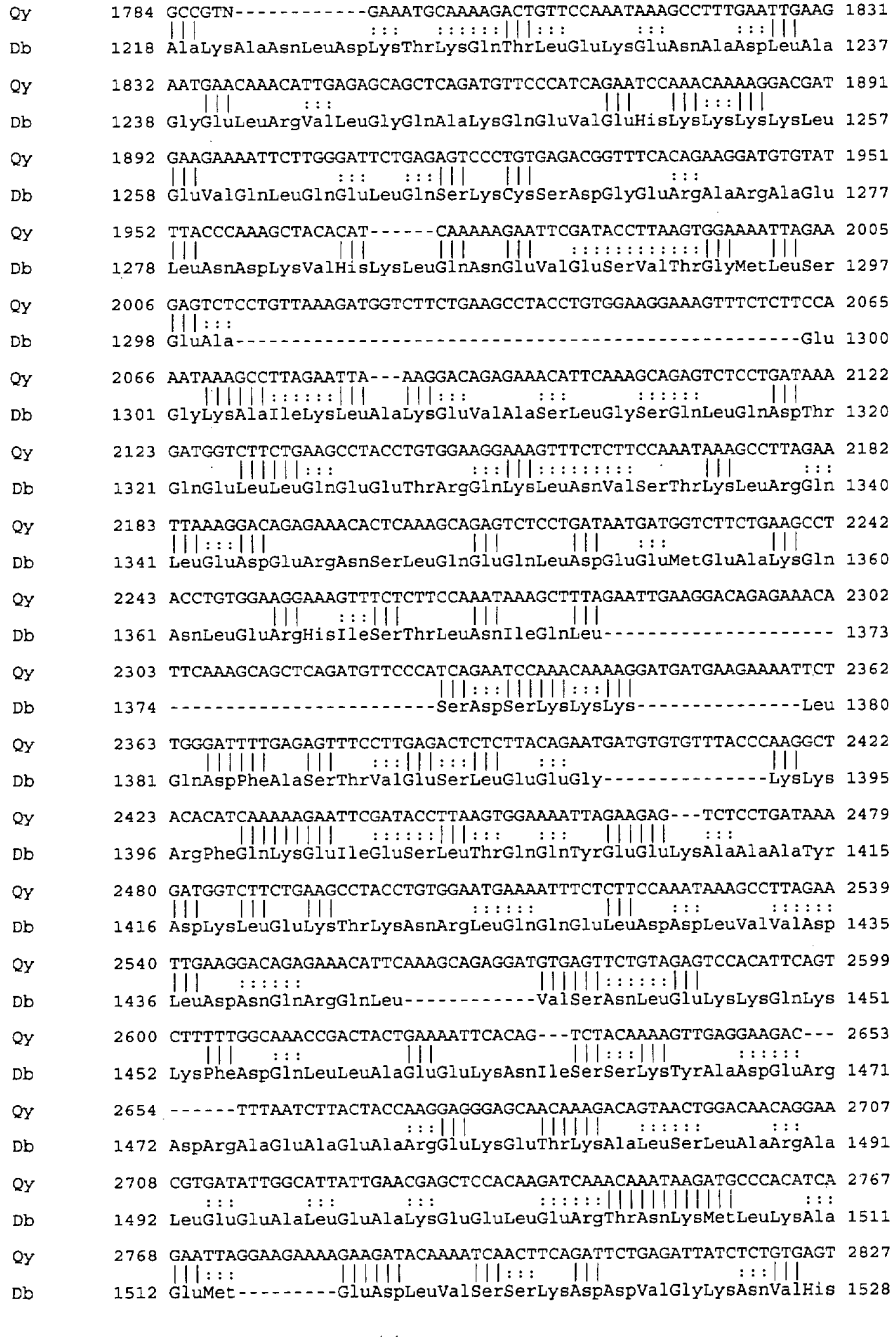
Alignment Scores:

Pred. No.:	4.97e-10	Length:	1972
Score:	315.00	Matches:	227
Percent Similarity:	37.99%	Conservative:	203
Best Local Similarity:	20.05%	Mismatches:	460
Query Match:	4.89%	Indels:	242
DB:	1	Gaps:	41

US-09-602-362E-26 (1-3673) x A41604 (1-1972)

```
Qy        677 GAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAG 736
              |||   |||   ||||||            |||     :::|||::::::
Db        866 GluArgGlnGlnLysAlaGluSerGlu------LeuGlnGluLeuGlnGlnLysHisThr 883

Qy        737 CAAACTGTGGAA------TTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAG 790
                  |||      |||         ||||||   :::         |||
```

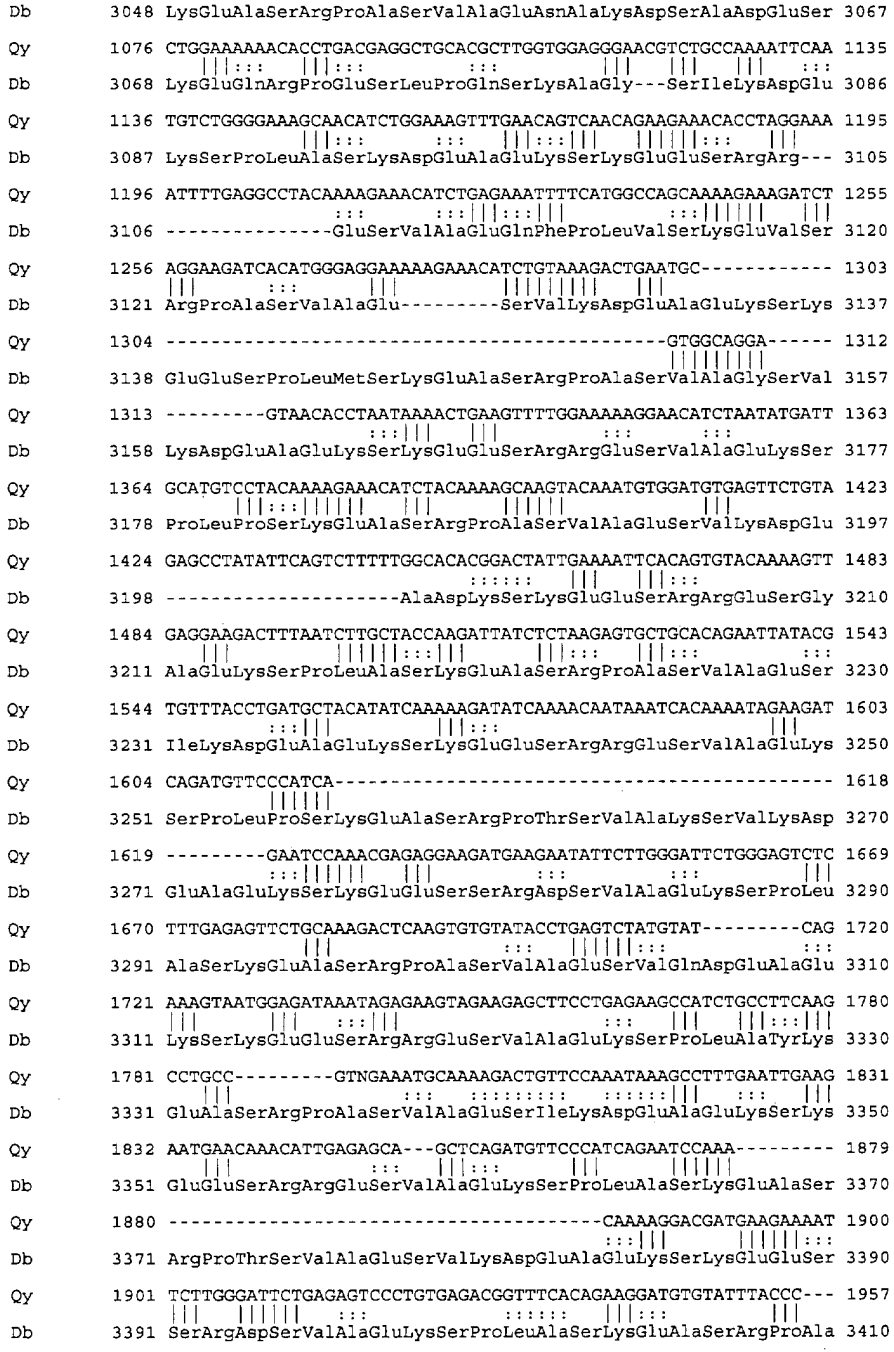
```
Db        884 GlnLeuSerGluGluLysAsnLeuLeuGlnGluGlnLeuGlnAlaGluThrGluLeuTyr 903

Qy        791 TCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATG 850
              ::::::         :::   :::   :::                |||:::      :::
Db        904 AlaGluAlaGluGluMetArgValArgLeuAlaAlaLysLysGlnGluLeuGluGluIle 923

Qy        851 CTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGA------ATAACTGCA 904
              |||   :::          :::      |||:::      |||      :::   |||
Db        924 LeuHisGluMetGluAlaArgLeuGluGluGluGluAspArgGlyGlnGlnLeuGlnAla 943

Qy        905 GAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACAT--- 961
              ||||||          |||:::   :::       :::   :::||||||   |||
Db        944 GluArgLysLysMetAlaGlnGlnMetLeuAspLeuGluGluGlnLeuGluGluGluGlu 963

Qy        962 ------------------------------------------------ATACGAAAATTA 973
                                                               |||:::||||||
Db        964 AlaAlaArgGlnLysLeuGlnLeuGluLysValThrAlaGluAlaLysIleLysLysLeu 983

Qy        974 CCT---------------AAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGA 1018
                              :::   ||||||
Db        984 GluAspAspIleLeuValMetAspAspGlnAsnAsnLys---------------------- 996

Qy       1019 ACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTG 1078
                  |||         |||   ||||||      |||         :::|||
Db        997 LeuSerLysGluArgLysLeuLeuGluGluArgIleSerAspLeuThrThrAsnLeuAla 1016

Qy       1079 GAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGT 1138
              |||:::      ::::::|||      |||   :::
Db       1017 GluGlu---GluGluLysAlaLysAsnLeuThrLys------------------------ 1027

Qy       1139 CTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATT 1198
                                  |||   |||           |||         ::::::
Db       1028 -----------LeuLysAsnLysHisGluSerMetIleSerGluLeuGluValArgLeu 1043

Qy       1199 TTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258
                 :::   |||      |||                  ::::::      |||
Db       1044 LysLysGluGluLysSerArgGlnGlu--------------LeuGluLysLeuLysArg 1058

Qy       1259 AAGATCACATGGGAGGAAAAAGAAACATCTGTAAAG---ACTGAATGCGTGGCAGGAGTA 1315
              ||:::      |||      :::      :::      :::      |||   :::
Db       1059 LysMetAspGlyGluAlaSerAspLeuHisGluGlnIleAlaAspLeuGlnAlaGlnIle 1078

Qy       1316 ACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACA 1375
                  |||   ::::::   :::|||                  |||
Db       1079 AlaGluLeuLysMetGlnLeuAlaLysLysGluGluGluLeuGlnAlaAlaLeuAlaArg 1098

Qy       1376 AAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTC 1435
                 |||      |||      |||
Db       1099 LeuGluAspGluThrSerGlnLysAsnAsn---------------------------- 1108

Qy       1436 AGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTT 1495
              :::|||      |||   :::|||      :::   ::::::||||||
Db       1109 AlaLeuLysLysIleArgGluLeuGlu---GlyHisIleSerAspLeuGlnGluAspLeu 1127

Qy       1496 AATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGAT 1555
              :::                  |||:::   ||||||:::|||
Db       1128 Asp-----------------SerGluArgAlaAlaArgAsnLysAlaGluLysGlnLys 1141

Qy       1556 GCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCA 1615
                           ::::::::::::   :::      ::::::||||||   :::
Db       1142 ArgAspLeuGlyGluGluLeuGluAlaLeuLysThrGluLeuGluAspThrLeuAspThr 1161

Qy       1616 TCAGAATCCAAACGAGAG------------GAAGATGAAGAATATTCTTGGGATTCTGGG 1663
              :::   :::::::::|||               :::::::|||   :::
Db       1162 ThrAlaThrGlnGlnGluLeuArgAlaLysArgGluGlnGluValThrValLeuLysLys 1181

Qy       1664 AGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAA 1723
              ::::|||   |||   :::      :::      :::   |||   |||   ||||||
Db       1182 AlaLeuAspGluGluThrArgSerHisGluAlaGlnValGlnGlu---MetArgGlnLys 1200

Qy       1724 GTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCT 1783
                    ::::::         ||||||||||||   |||:::         ||||||
Db       1201 HisThrGlnVal---------ValGluGluLeuThrGluGlnLeuGluGlnPheLysArg 1217
```


Qy 1784 GCCGTN----------GAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAG 1831
Db 1218 AlaLysAlaAsnLeuAspLysThrLysGlnThrLeuGluLysGluAsnAlaAspLeuAla 1237
Qy 1832 AATGAACAAACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGAT 1891
Db 1238 GlyGluLeuArgValLeuGlyGlnAlaLysGlnGluValGluHisLysLysLysLysLeu 1257
Qy 1892 GAAGAAAATTCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTAT 1951
Db 1258 GluValGlnLeuGlnGluLeuGlnSerLysCysSerAspGlyGluArgAlaArgAlaGlu 1277
Qy 1952 TTACCCAAAGCTACACAT------CAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAA 2005
Db 1278 LeuAsnAspLysValHisLysLeuGlnAsnGluValGluSerValThrGlyMetLeuSer 1297
Qy 2006 GAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCA 2065
Db 1298 GluAla----------Glu 1300
Qy 2066 AATAAAGCCTTAGAATTA---AAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAA 2122
Db 1301 GlyLysAlaIleLysLeuAlaLysGluValAlaSerLeuGlySerGlnLeuGlnAspThr 1320
Qy 2123 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2182
Db 1321 GlnGluLeuLeuGlnGluGluThrArgGlnLysLeuAsnValSerThrLysLeuArgGln 1340
Qy 2183 TTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCT 2242
Db 1341 LeuGluAspGluArgAsnSerLeuGlnGluGlnLeuAspGluGluMetGluAlaLysGln 1360
Qy 2243 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACA 2302
Db 1361 AsnLeuGluArgHisIleSerThrLeuAsnIleGlnLeu---------- 1373
Qy 2303 TTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCT 2362
Db 1374 ----------SerAspSerLysLysLys----------Leu 1380
Qy 2363 TGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCT 2422
Db 1381 GlnAspPheAlaSerThrValGluSerLeuGluGluGly----------LysLys 1395
Qy 2423 ACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG---TCTCCTGATAAA 2479
Db 1396 ArgPheGlnLysGluIleGluSerLeuThrGlnGlnTyrGluGluLysAlaAlaAlaTyr 1415
Qy 2480 GATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAA 2539
Db 1416 AspLysLeuGluLysThrLysAsnArgLeuGlnGlnGluLeuAspAspLeuValValAsp 1435
Qy 2540 TTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGT 2599
Db 1436 LeuAspAsnGlnArgGlnLeu----------ValSerAsnLeuGluLysLysGlnLys 1451
Qy 2600 CTTTTTGGCAAACCGACTACTGAAAATTCACAG---TCTACAAAAGTTGAGGAAGAC--- 2653
Db 1452 LysPheAspGlnLeuLeuAlaGluGluLysAsnIleSerSerLysTyrAlaAspGluArg 1471
Qy 2654 ------TTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAA 2707
Db 1472 AspArgAlaGluAlaGluAlaArgGluLysGluThrLysAlaLeuSerLeuAlaArgAla 1491
Qy 2708 CGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA 2767
Db 1492 LeuGluGluAlaLeuGluAlaLysGluGluLeuGluArgThrAsnLysMetLeuLysAla 1511
Qy 2768 GAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGTGAGT 2827
Db 1512 GluMet---------GluAspLeuValSerSerLysAspAspValGlyLysAsnValHis 1528
Qy 2828 GATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACA--- 2884
Db 1529 GluLeuGluLysSerLysArgAlaLeuGlu---ThrGlnMetGluGluMetLysThrGln 1547
Qy 2884 ---------- 2884
Db 1548 LeuGluGluLeuGluAspGluLeuGlnAlaThrGluAspAlaLysLeuArgLeuGluVal 1567
Qy 2885 ----------ACAAATGGCAAAATAGAAGAGTCT 2908
Db 1568 AsnMetGlnAlaLeuLysValGlnPheGluArgAspLeuGlnAlaArgAspGluGlnAsn 1587
Qy 2909 CCTGAAAAGCCTTCTCACTTT----------GAGCCTGCCACTGAAATGCAAAAC 2953
Db 1588 GluGluLysArgArgGlnLeuGlnArgGlnLeuHisGluTyrGluThrGluLeuGluAsp 1607
Qy 2954 TCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACT 3013
Db 1608 GluArgLysGlnArgAlaLeuAlaAlaAlaAlaLysLysLysLeuGluGlyAspLeuLys 1627
Qy 3014 ACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAA 3073
Db 1628 AspLeuGluLeuGlnAlaAsp---------SerAlaIleLysGlyArgGlu--------- 1641
Qy 3074 GATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTA 3133
Db 1642 GluAlaIleLysGlnLeuLeuLysGlnAlaGlnMetLysAsp----------Phe 1657
Qy 3134 CAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTA----------GAG 3178
Db 1658 GlnArgGluLeuGluAspAlaArgAlaSerArgAspGluIlePheAlaThrAlaLysGlu 1677
Qy 3179 AACCAAAAA------GCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAAT 3232
Db 1678 AsnGluLysLysAlaLysSerLeuGluAlaAspLeuMetGlnLeuGlnGluAspLeuAla 1697
Qy 3233 CAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAA--- 3289
Db 1698 AlaAlaGluArgAlaArgLysGlnAlaAspLeuGluLysGluGluLeuAlaGluGluLeu 1717
Qy 3290 ----------GAGCAACTTAGGAAAAAGTTAGAAGTGAAA 3319
Db 1718 AlaSerSerLeuSerGlyArgAsnAlaLeuGlnAspGluLysArgArgLeuGluAlaArg 1737
Qy 3320 ---CACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTA------ 3370
Db 1738 IleAlaGlnLeuGluGluGluLeuGluGluGluGlnGlyAsnMetGluAlaMetSerAsp 1757
Qy 3371 ----------ACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAAT 3415
Db 1758 ArgValArgLysAlaThrGlnGlnAlaGluGlnLeuSerAsnGluLeuAlaThrGluArg 1777
Qy 3416 GATCTCTTTCATGAAAAT---------- 3433
Db 1778 SerThrAlaGlnLysAsnGluSerAlaArgGlnGlnLeuGluArgGlnAsnLysGluLeu 1797
Qy 3434 ----------TGCATGTTGAAAAAGGAAATTGCC 3457
Db 1798 LysSerLysLeuGlnGluMetGluGlyAlaValLysSerLysPheLysSerThrIleAla 1817
Qy 3458 ATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATAC 3517
Db 1818 AlaLeuGluAlaLysIleAlaGlnLeuGluGluGlnValGluGlnGluAlaArgGluLys 1837
Qy 3518 TTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTG 3577
Db 1838 GlnAlaAlaAlaLysAlaLeuLysGlnArgAspLysLysLeuLysGluMetLeu---Leu 1856
Qy 3578 AAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACG 3637
Db 1857 GlnValGluAspGluArgLysMetAlaGluGlnTyrLysGluGlnAlaGluLysGlyAsn 1876
Qy 3638 GCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673

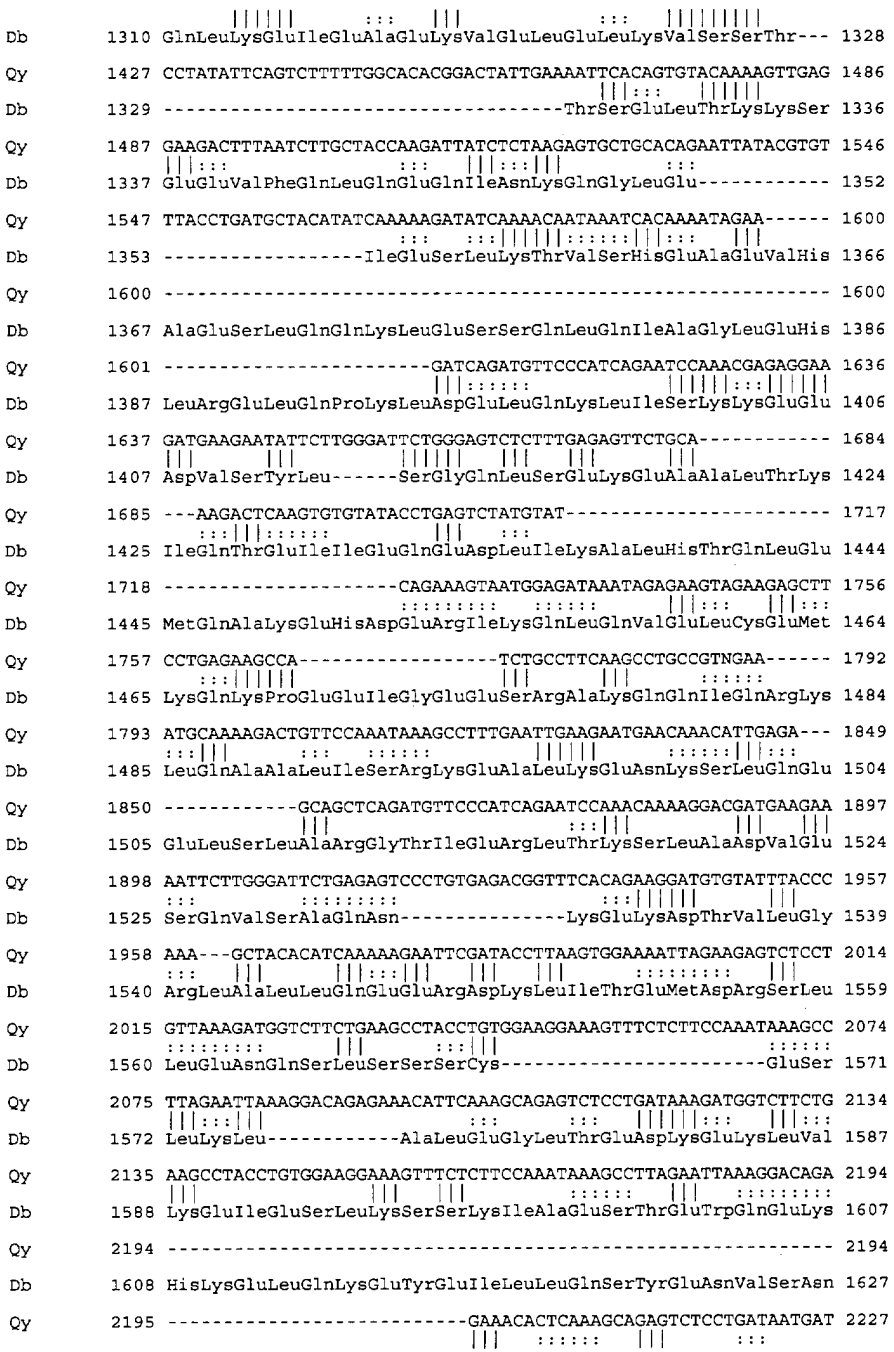
```
Db       1877 AlaLysValLysGlnLeuLysArgGlnLeuGluGlu 1888

RESULT 10
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34418
R;Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid F12F3.
A;Reference number: Z21521
A;Accession: T34418
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3488 <FUL>
A;Cross-references: EMBL:U80022; PIDN:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
A;Experimental source: strain Bristol N2; clone F12F3
C;Genetics:
A;Gene: CESP:F12F3.3
A;Map position: 5
A;Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Alignment Scores:
Pred. No.:                5.3e-10        Length:         3488
Score:                    315.00         Matches:        225
Percent Similarity:       37.33%         Conservative:   161
Best Local Similarity:    21.76%         Mismatches:     424
Query Match:              4.89%          Indels:         224
DB:                       2              Gaps:           43

US-09-602-362E-26 (1-3673) x T34418 (1-3488)

Qy        974 CCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCT 1033
Db        486 ProSerSerSerLysGluSerSerPro------ProThrAspGlyLysLysLysLysGln 503

Qy       1034 GCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTG--------------- 1078
Db        504 IleProLysAlaLeuPheIleProAspGluIleSerSerArgPheGlyAspProSerThr 523

Qy       1079 ----------------------------------------GAAAAAACACCTGACGAG 1096
Db        524 MetHisSerGluThrAsnIleThrThrThrIleArgGlyArgGluGlySerAlaAspAla 543

Qy       1097 GCTGCACGCTTGGTGGAGGGAACGTCTGCC-----------AAAATTCAATGTCTGGGG 1144
Db        544 LysThrProLeuValGluProLeuSerAlaSerValSerMetLysValPheThrLeuVal 563

Qy       1145 AAAGCAACATCTGGAAAGTTTGAACAGTCA---------ACAGAAGAAACACCT------ 1189
Db        564 GluSerAlaLysGluLysAlaGluPheSerPheLysArgArgSerGluThrProAspAsp 583

Qy       1190 ------------AGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGG 1237
Db        584 LysSerArgLysLysGluGlyLeuProProAlaLysLysSerGluLysLysAspGluVal 603

Qy       1238 CCAGCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACT 1297
Db        604 ThrAlaGluLysGlnSerThrGluAlaLeuIleGluSerLysLysLysGluValAspGlu 623

Qy       1298 GAATGCGTGGCAGGAGTAACACCT---------AATAAAACTGAAGTTTTG--------- 1339
Db        624 SerLysIleSerGluGlnGlnProSerAspLysAsnLysSerGluValValGlyValPro 643

Qy       1340 ----------------------GAAAAAGGAACATCTAATATGATTGCATGTCCTACA 1375
Db        644 GluLysAlaAlaGlyProGluThrLysLysAspValSerGluIleGluGluValProLys 663

Qy       1376 AAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTC 1435
Db        664 LysLysThrIleLysLysLysThrGluLysSerAsp---SerSerIleSerGlnLysSer 682

Qy       1436 AGTCTTTTTGGCACACGGACTATTGAAAATTCA--------------------CAGTGT 1474
Db        683 AsnValLeuLysProAlaAspAspAspLysSerLysSerAspAspValThrAspLysSer 702

Qy       1475 ACAAAAGTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAG 1534
Db        703 LysLysThrThrGluAsp---------GlnThrLysValAlaThrAspSerLysLeuGlu 719

Qy       1535 AATTATACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAA 1594
Db        720 Lys---------AlaAlaAspThrThrLysGlnIleGluThrGluThrVal--------- 733

Qy       1595 ATAGAAGATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGG 1654
Db        734 ValAspAsp--------------LysSerLysLysLysValLeuLysLysLysThrGlu 748

Qy       1655 GATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATA------------ 1702
Db        749 LysSerAspSerPheIleSerGlnLysSerGluThrProProValValGluProThrLys 768

Qy       1703 ---CCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCT 1759
Db        769 ProAlaGluSerGluAlaGlnLysIleAlaGluValAsnLysAlaLysLysGlnLysGlu 788

Qy       1760 GAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCC 1819
Db        789 ValAspAspAsnLeuLysArgGluAlaGluValAlaAlaLysLysIleAlaAspGluLys 808

Qy       1820 TTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAA 1879
Db        809 LeuLysIleGluAlaGluAlaAsnIleLysLysThrAlaGluValGluAlaAlaLysLys 828

Qy       1880 CAAAAGGAC---GATGAAGAAAATTCTTGGGATTCTGAG---------------AGTCCC 1921
Db        829 GlnLysGluLysAspGluGlnLeuLysLeuGluThrGluValValSerLysLysSerAla 848

Qy       1922 TGTGAGACGGTTTCA---CAGAAGGATGTGTATTTACCCAAAGCTACACAT--------- 1969
Db        849 AlaGluLysLeuGluLeuGluLysGlnAlaGlnIleLysLysAlaAlaGluAlaAspAla 868

Qy       1970 ---------CAAAAAGAATTCGAT---------------------------------- 1984
Db        869 ValLysLysGlnLysGluLeuAsnGluLysAsnLysLeuGluAlaAlaLysLysSerAla 888

Qy       1985 --------------------------------ACCTTAAGTGGAAAATTAGAAGAGTCT 2011
Db        889 AlaAspLysLeuLysLeuGluGluGluSerAlaAlaLysSerLysLysValSerGluGlu 908

Qy       2012 CCTGTTAAAGATGGTCTTCTGAAGCCTACC---TGTGGAAGGAAAGTT------------ 2056
Db        909 SerValLysPheGlyGluGluLysLysThrLysAlaGlyGluLysThrValGlnValGlu 928

Qy       2057 TCTCTTCCA------AATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAAAGCA--- 2107
Db        929 SerGluProThrSerLysLysThrIleAspThrLysAspValGlyAlaThrGluProAla 948

Qy       2108 ---GAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTT 2164
Db        949 AspGluThrProLysLysLysIleIleLysLysLysThrGluLysSerAspSerSerIle 968

Qy       2165 CCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAAT 2224
Db        969 SerGlnLysSer------AlaThrAspSerGluLysVal---SerLysGlnLysGluGln 985

Qy       2225 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGAA 2284
Db        986 AspGluProThrLysProAlaValSerGluThrGlnMetValThrGluAlaAspLysSer 1005

Qy       2285 TTGAAGGACAGAGAAACA---------TTCAAAGCAGCTCAGATGTTCCCATCAGAATCC 2335
Db       1006 LysLysGlnLysGluThrAspGluLysLeuLysLeuAspAlaGluIleAlaAlaLysThr 1025
```

```
Qy       2336 AAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTA 2395
Db       1026 LysGlnGluAlaAspGluLysSerLysLeuAspAlaGlnGluLysIleLysLysValSer 1045

Qy       2396 CAGAATGATGTGTGTTTACCCAAG---------------------------------- 2419
Db       1046 GluAspAspAlaAlaArgLysGluLysGluLeuAsnAspLysLeuLysLeuGluSerGlu 1065

Qy       2420 ---GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG--------- 2467
Db       1066 IleAlaThrLysLysAlaSerAlaAspLysLeu------LysLeuGluGluGlnAlaGln 1083

Qy       2468 ----------------------------------TCTCCTGATAAAGATGGTCTTCTG 2491
Db       1084 AlaLysLysAlaAlaGluValGluAlaAlaLysLysGlnLysGluLysAspGluGlnLeu 1103

Qy       2492 AAG---------CCTACCTGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTG 2542
Db       1104 LysLeuAspThrGluAlaAlaSerLysLysAlaAlaAlaGluLysLeuGluLeuGlu--- 1122

Qy       2543 AAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTT 2602
Db       1123 LysGlnAlaGlnIleLysLysAlaAlaGlyAlaAspAlaValLysLysGlnLysGluLeu 1142

Qy       2603 TTTGGCAAACCGACTACTGAA---AATTCACAGTCTACA------------AAAGTTGAG 2647
Db       1143 AspGluLysAsnLysLeuGluAlaAsnLysLysSerAlaAlaGlyLysLeuLysIleGlu 1162

Qy       2648 GAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAGGAA 2707
Db       1163 GluGluSerAlaAlaLysSerLysGln------------------------------- 1171

Qy       2708 CGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACATCA 2767
Db       1172 ---------ThrValGluGluGlnAlaLysLeuAspAlaGlnThrLysAlaLysThrAla 1188

Qy       2768 GAATTAGGAAGAAAA------GAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCT 2821
Db       1189 GluLysGlnThrLysLeuGluLysAspGluLysSerThrLysGluSerGluSerLysGlu 1208

Qy       2822 GTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAG 2881
Db       1209 ThrValAspGluLysProLysLysLysValLeuLysLysLysThrGluLysSerAspSer 1228

Qy       2882 ACAACAAATGGCAAAATAGAAGAGTCTCCT---------GAAAAGCCTTCTCACTTTGAG 2932
Db       1229 SerIleSerGlnLysSerGluThrSerLysThrValValGluSerAlaGlyProSerGlu 1248

Qy       2933 CCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAA 2992
Db       1249 SerGluThrGlnLysValAlaAspAlaAlaArgLysGlnLysGluThrAspGluLysGln 1268

Qy       2993 ACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAA 3052
Db       1269 LysLeuGluAlaGlu---IleThrAlaLysLysSerAlaAspGluLysSerLysLeuGlu 1287

Qy       3053 AGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATG 3112
Db       1288 AlaGluSerLysLeuLysLys-----------------AlaAlaGluValGluAlaAla 1301

Qy       3113 AAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAG 3172
Db       1302 LysLys--------------GlnLysGluLysAspGluGlnLeuLysLeuAspThrGlu 1316

Qy       3173 TTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAAT 3232
Db       1317 AlaAlaSerLysLysAlaAlaAlaGluLys-----------LeuGluLeuGluLysGln 1332

Qy       3233 CAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAAGAG 3292
Db       1333 SerHisIleLysLysAlaAlaGluValAspAlaValLysLysGlnLysGluLeuGluGlu 1352

Qy       3293 CAACTTAGG---------------------------------AAAAAGTTAGAAGTG 3316
```

```
Db       1353 LysGlnArgLeuGluSerGluAlaAlaThrLysLysAlaAspAlaGluLysLeuLysLeu 1372

Qy       3317 AAACACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGT 3376
Db       1373 GluGluGlnLysLysLysAlaAlaGluIleAlaLeuIleGluIleGlnLysGluGlnGlu 1392

Qy       3377 AATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGC 3436
Db       1393 LysLeuAlaGlnGluGlnSerArgLeuGluAspGlu------------------------ 1404

Qy       3437 ATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACAC 3496
Db       1405 ---AlaLysLysSerAlaGluLysGlnLysLeuGluSerGluThrLysSerLysGlnThr 1423

Qy       3497 CAGGTGAAGGAAAATAAATACTTTGAGGAC---------ATTAAGATTTTACAAGAAAAG 3547
Db       1424 GluGluAlaProLysGluSerValAspGluLysProLysLysLysValLeuLysLysLys 1443

Qy       3548 AATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACA 3589
Db       1444 ThrGluLysSerAspSerSerIleSerGlnLysSerLysSer 1457

RESULT 11
T13564
microtubule-associated protein homolog - fruit fly (Drosophila melanogaster)
N;Alternate names: hypothetical protein EG:49E4.1
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13564
R;Spanos, L.; Papagiannakis, G.; Siden-Kiamos, I.; Louis, C.
submitted to the EMBL Data Library, April 1999
A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A;Reference number: Z17689
A;Accession: T13564
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5327 <SPA>
A;Cross-references: EMBL:AL031128; PIDN:CAA20006.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0025392
A;Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1
A;Note: EG:49E4.1
C;Superfamily: Drosophila 576K microtubule-associated protein homolog

Alignment Scores:
Pred. No.:                7.65e-10        Length:          5327
Score:                    312.50          Matches:         228
Percent Similarity:       34.80%          Conservative:    167
Best Local Similarity:    20.09%          Mismatches:      485
Query Match:              4.85%           Indels:          255
DB:                       2               Gaps:            39

US-09-602-362E-26 (1-3673) x T13564 (1-5327)

Qy        794 AAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTT 853
Db       2984 LysThrThrSerArgArgValSerValAlaAspSerIleLysAspGluLysSerLeuLeu 3003

Qy        854 CTTCAGCAAAATGTT---------GACGTCTTTGCTGAAGACATACATGGAATAACTGCA 904
Db       3004 ValSerGlnGluAlaSerArgProGluSerGluAlaGluSerLeuLysAspAlaAlaAla 3023

Qy        905 GAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATA 964
Db       3024 ProSerGlnGluThrSerArg------------------------------------- 3030

Qy        965 CGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCT 1024
Db       3031 ---------ProGluSerValThrGluSerValLysAspGlyLysSerProValAlaSer 3047

Qy       1025 GATGAGGCT---------GCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTG 1075
```


Db 3048 LysGluAlaSerArgProAlaSerValAlaGluAsnAlaLysAspSerAlaAspGluSer 3067
Qy 1076 CTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAA 1135
Db 3068 LysGluGlnArgProGluSerLeuProGlnSerLysAlaGly---SerIleLysAspGlu 3086
Qy 1136 TGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAA 1195
Db 3087 LysSerProLeuAlaSerLysAspGluAlaGluLysSerLysGluGluSerArgArg--- 3105
Qy 1196 ATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCT 1255
Db 3106 --------------GluSerValAlaGluGlnPheProLeuValSerLysGluValSer 3120
Qy 1256 AGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGC------------ 1303
Db 3121 ArgProAlaSerValAlaGlu---------SerValLysAspGluAlaGluLysSerLys 3137
Qy 1304 --GTGGCAGGA------ 1312
Db 3138 GluGluSerProLeuMetSerLysGluAlaSerArgProAlaSerValAlaGlySerVal 3157
Qy 1313 ---------GTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATT 1363
Db 3158 LysAspGluAlaGluLysSerLysGluGluSerArgArgGluSerValAlaGluLysSer 3177
Qy 1364 GCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTA 1423
Db 3178 ProLeuProSerLysGluAlaSerArgProAlaSerValAlaGluSerValLysAspGlu 3197
Qy 1424 GAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTT 1483
Db 3198 --------------------AlaAspLysSerLysGluGluSerArgArgGluSerGly 3210
Qy 1484 GAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACG 1543
Db 3211 AlaGluLysSerProLeuAlaSerLysGluAlaSerArgProAlaSerValAlaGluSer 3230
Qy 1544 TGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGAT 1603
Db 3231 IleLysAspGluAlaGluLysSerLysGluGluSerArgArgGluSerValAlaGluLys 3250
Qy 1604 CAGATGTTCCCATCA-- 1618
Db 3251 SerProLeuProSerLysGluAlaSerArgProThrSerValAlaLysSerValLysAsp 3270
Qy 1619 ---------GAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTC 1669
Db 3271 GluAlaGluLysSerLysGluGluSerSerArgAspSerValAlaGluLysSerProLeu 3290
Qy 1670 TTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTAT---------CAG 1720
Db 3291 AlaSerLysGluAlaSerArgProAlaSerValAlaGluSerValGlnAspGluAlaGlu 3310
Qy 1721 AAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAG 1780
Db 3311 LysSerLysGluGluSerArgArgGluSerValAlaGluLysSerProLeuAlaTyrLys 3330
Qy 1781 CCTGCC---------GTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAG 1831
Db 3331 GluAlaSerArgProAlaSerValAlaGluSerIleLysAspGluAlaGluLysSerLys 3350
Qy 1832 AATGAACAAACATTGAGAGCA---GCTCAGATGTTCCCATCAGAATCCAAA--------- 1879
Db 3351 GluGluSerArgArgGluSerValAlaGluLysSerProLeuAlaSerLysGluAlaSer 3370
Qy 1880 ------------------------------------CAAAAGGACGATGAAGAAAAT 1900
Db 3371 ArgProThrSerValAlaGluSerValLysAspGluAlaGluLysSerLysGluGluSer 3390
Qy 1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCC--- 1957
Db 3391 SerArgAspSerValAlaGluLysSerProLeuAlaSerLysGluAlaSerArgProAla 3410
Qy 1958 ----------------------AAAGCTACACATCAAAAAGAATTCGATACCTTAAGT 1993
Db 3411 SerValAlaGluSerValGlnAspGluAlaGluLysSerLysGluGluSerArgArgGlu 3430
Qy 1994 GGAAAATTAGAAGAGTCTCCT-- 2014
Db 3431 SerValAlaGluLysSerProLeuAlaSerLysGluAlaSerArgProAlaSerValAla 3450
Qy 2015 ------GTTAAAGATGGTCTTCTGAAG------------------------------- 2035
Db 3451 GluSerValLysAspAspAlaGluLysSerLysGluGluSerArgArgGluSerValAla 3470
Qy 2036 ---------CCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA---TTA 2083
Db 3471 GluLysSerProLeuAlaSerLysGluAlaSerArgProAlaSerValAlaGluSerVal 3490
Qy 2084 AAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAG------ 2137
Db 3491 LysAspGluAlaGluLysSerLysGluGluSerArgArgGluSerValAlaGluLysSer 3510
Qy 2138 CCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTA------------ 2185
Db 3511 ProLeuProSerLysGluAlaSerArgProThrSerValAlaGluSerValLysAspGlu 3530
Qy 2186 ------AAGGACAGAGAAACACTCAAAGCAGAGTCT---CCTGATAATGATGGTCTTCTG 2236
Db 3531 AlaGluLysSerLysGluGluSerArgArgGluSerValAlaGluLysSerSerLeuAla 3550
Qy 2237 AAGCCTACCTGTGGAAGGAAAGTTTCTCTT---------------------CCAAATAAA 2275
Db 3551 SerLysGluAlaSerArgProAlaSerValAlaGluSerValLysAspGluAlaGluLys 3570
Qy 2276 GCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCC 2335
Db 3571 SerLysGluGluSerArgArgGluSer------ValAlaGluLysSerProLeuAlaSer 3588
Qy 2336 AAACAA--------------------------------AAGGATGATGAAGAAAATTCT 2362
Db 3589 LysGluAlaSerArgProAlaSerValAlaGluSerValLysAspGluAlaGluLysSer 3608
Qy 2363 TGGGATTTT---------GAGAGTTTCCTTGAG------ACTCTCTTACAGAATGATGTG 2407
Db 3609 LysGluValSerArgArgGluSerValAlaGluLysSerProLeuProSerLysGluAla 3628
Qy 2408 TGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAG 2467
Db 3629 SerArgProThrSerValAlaGluSerValLysAspGluAlaAspLysSerLysGluGlu 3648
Qy 2468 TCTCCTGATAAAGATGGTCTTCTGAAG---CCTACCTGTGGAATGAAAATTTCTCTTCCA 2524
Db 3649 SerArgArgGluSerGlyAlaGluLysSerProLeuAlaSerMetGluAlaSerArgPro 3668
Qy 2525 AATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGTA 2584
Db 3669 ThrSerValAlaGluSerValLysAspGluThrGluLysSerLysGluGluSerArgArg 3688
Qy 2585 GAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGTT 2644
Db 3689 GluSerValThrGluLysSerProLeuProSerLysGluAlaSerArgProThrSerVal 3708
Qy 2645 GAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACAG 2704
Db 3709 AlaGlu------------SerValLysAspGluAlaGluLys------SerLysGluGlu 3722
Qy 2705 GAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCACA 2764
Db 3723 SerArgArgGluSerValAlaGluLysSerPro---------------------------- 3733
Qy 2765 TCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATC------ 2818
Db 3734 ------LeuAlaSerLysGluSerSerArgProAlaSerValAlaGluSerIleLysAsp 3751

```
Qy       2819 ---------TCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAA 2869
Db       3752 GluAlaGluGlyThrLysGlnGluSerArgArgGluSerMetProGluSerGlyLysAla 3771

Qy       2870 AAAGAAATAAAGACAACAAATGGCAAAATA---GAAGAGTCTCCTGAAAAGCCTTCTCAC 2926
Db       3772 GluSerIleLysGlyAspGlnSerSerLeuAlaSerLysGluThrSerArgProAspSer 3791

Qy       2927 TTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAAT 2986
Db       3792 ValValGluSerValLysAspGluThrGluLysProGluGlySerAlaIleAspLysSer 3811

Qy       2987 AAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCT 3046
Db       3812 GlnValAlaSerArgProGluSerValAlaValSer---------------------- 3823

Qy       3047 TGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAA 3106
Db       3824 --------------------AlaLysAspGluLysSerProLeuHisSerArgProGlu 3836

Qy       3107 CAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAA 3166
Db       3837 SerValAlaAspLysSerProAspAlaSerLysGluAlaSerArgSerLeuSerValAla 3856

Qy       3167 TCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCT 3226
Db       3857 GluThrAlaSerSerProIleGluGluGlyProArgSerIleAlaAspLeuSerLeuPro 3876

Qy       3227 TTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC 3286
Db       3877 LeuAsnLeuThrGlyGluAlaLysGlyLysLeuProThrLeuSerSerProIleAspVal 3896

Qy       3287 GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTT------------------ 3328
Db       3897 AlaGlu------GlyAspPheLeuGluValLysAlaGluSerSerProArgProAlaVal 3914

Qy       3328 ---------------------------------------------------------- 3328
Db       3915 LeuSerLysProAlaGluPheSerGlnProAspThrGlyHisThrAlaSerThrProVal 3934

Qy       3329 ---GAACAGACTCTCAGAATACAAGATATAGAATTG------AAAAGTGTAACAAGTAAT 3379
Db       3935 AspGluAlaSerProValLeuGluGluIleGluValValGluGlnHisThrThrSerGly 3954

Qy       3380 TTGAATCAGGTTTCTCACACTCATGAAAGTGAA------------AATGATCTCTTTCAT 3427
Db       3955 ValGlyAlaThrGlyAlaThrAlaGluThrAspLeuLeuAspLeuThrGluThrLysSer 3974

Qy       3428 GAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAA 3487
Db       3975 GluThrValThrLysGlnSerGluThrThrLeuPheGluThrLeuThrSerLysValGlu 3994

Qy       3488 CATCAACACCAGGTGAAGGAAAAT------AAATACTTTGAGGACATTAAGATTTTACAA 3541
Db       3995 SerLysValGluValLeuGluSerSerValLysGlnValGluGluLysValGlnThrSer 4014

Qy       3542 GAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGG 3601
Db       4015 ValLysGlnAlaGluThrThrValThrAspSerLeuGluGln------LeuThrLysLys 4032

Qy       3602 GCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAAC 3646
Db       4033 SerSerGluGlnLeuThrGluIleLysSerValLeuAspThrAsn 4047

RESULT 12
A56539
giantin - human
N;Alternate names: macrogolgin
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 26-Jan-1996 #text_change 10-Dec-1999
C;Accession: A56539; S37536
R;Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.
Mol. Cell. Biol. 14, 2564-2576, 1994
A;Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein (g
A;Reference number: A56539; MUID:94187728; PMID:7511208
A;Accession: A56539
A;Molecule type: mRNA
A;Residues: 1-3259 <SEE>
A;Cross-references: EMBL:X75304; NID:g405714; PIDN:CAA53052.1; PID:g405715
C;Genetics:
A;Gene: GDB:GOLGB1; GCP; GCP371
A;Cross-references: GDB:454958
A;Map position: 3q13.31-3q13.31
C;Superfamily: giantin
C;Keywords: coiled coil; Golgi apparatus; transmembrane protein
F;3238-3254/Domain: transmembrane #status predicted <TMN>

Alignment Scores:
Pred. No.:                8.77e-10        Length:          3259
Score:                    311.00          Matches:         235
Percent Similarity:       36.70%          Conservative:    212
Best Local Similarity:    19.29%          Mismatches:      425
Query Match:              4.82%           Indels:          346
DB:                       1               Gaps:            48

US-09-602-362E-26 (1-3673) x A56539 (1-3259)

Qy        680 GTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAA 739
Db       1101 LeuGlnAspLysThrAsnGlnIleAspLeuLeuGlnAlaGluIleSerGluAsnGlnAla 1120

Qy        740 ACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAG--------- 790
Db       1121 IleIleGlnLysLeuIleThrSerAsnThrAspAlaSerAspGlyAspSerValAlaLeu 1140

Qy        791 TCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATG 850
Db       1141 ValLysGluThrValValIleSerProProCysThrGlySerSerGluHisTrpLysPro 1160

Qy        851 CTTCTTCAGCAAAATGTT---------GACGTCTTTGCTGAAGACATACATGGAATAACT 901
Db       1161 GluLeuGluGluLysIleLeuAlaLeuGluLysGluLysGluGlnLeuGlnLysLysLeu 1180

Qy        902 GCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACAT 961
Db       1181 GlnGluAlaLeuThrSerArgLysAlaIleLeuLysLysAlaGlnGluLysGluArgHis 1200

Qy        962 ATACGA---------AAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCT 1012
Db       1201 LeuArgGluGluLeuLysGlnGlnLysAspAspTyrAsnArgLeuGlnGluGlnPhe--- 1219

Qy       1013 ACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGC 1072
Db       1220 -----------AspGluGlnSerLysGluAsnGluAsnIleGlyAspGlnLeuArgGln 1235

Qy       1073 TTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATT 1132
Db       1236 LeuGlnIleGlnValArgGluSerIleAspGlyLysLeuProSerThrAspGlnGlnGlu 1255

Qy       1133 CAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGG 1192
Db       1256 SerCys------SerSerThrProGlyLeuGluGluProLeuPheLysAlaThrGluGln 1273

Qy       1193 AAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGA 1252
Db       1274 HisHisThrGlnProValLeuGluSerAsnLeuCysProAspTrpProSerHisSerGlu 1293

Qy       1253 TCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGA 1312
Db       1294 Asp---------------------------------------------------------- 1294

Qy       1313 GTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCT---AATATGATTGCATGT 1369
Db       1295 --------------AlaSerAlaLeuGlnGlyGlyThrSerValAlaGlnIleLysAla 1309

Qy       1370 CCTACAAAAGAA---ACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAG 1426
```


Db 1310 GlnLeuLysGluIleGluAlaGluLysValGluLeuGluLeuLysValSerSerThr--- 1328
Qy 1427 CCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAG 1486
Db 1329 ----------ThrSerGluLeuThrLysLysSer 1336
Qy 1487 GAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGT 1546
Db 1337 GluGluValPheGlnLeuGlnGluGlnIleAsnLysGlnGlyLeuGlu---------- 1352
Qy 1547 TTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA------ 1600
Db 1353 ----------IleGluSerLeuLysThrValSerHisGluAlaGluValHis 1366
Qy 1600 ---------- 1600
Db 1367 AlaGluSerLeuGlnGlnLysLeuGluSerSerGlnLeuGlnIleAlaGlyLeuGluHis 1386
Qy 1601 ----------GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAA 1636
Db 1387 LeuArgGluLeuGlnProLysLeuAspGluLeuGlnLysLeuIleSerLysLysGluGlu 1406
Qy 1637 GATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCA---------- 1684
Db 1407 AspValSerTyrLeu------SerGlyGlnLeuSerGluLysGluAlaAlaLeuThrLys 1424
Qy 1685 ---AAGACTCAAGTGTGTATACCTGAGTCTATGTAT---------- 1717
Db 1425 IleGlnThrGluIleIleGluGlnGluAspLeuIleLysAlaLeuHisThrGlnLeuGlu 1444
Qy 1718 ----------CAGAAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTT 1756
Db 1445 MetGlnAlaLysGluHisAspGluArgIleLysGlnLeuGlnValGluLeuCysGluMet 1464
Qy 1757 CCTGAGAAGCCA----------TCTGCCTTCAAGCCTGCCGTNGAA------ 1792
Db 1465 LysGlnLysProGluGluIleGlyGluGluSerArgAlaLysGlnGlnIleGlnArgLys 1484
Qy 1793 ATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGA--- 1849
Db 1485 LeuGlnAlaAlaLeuIleSerArgLysGluAlaLeuLysGluAsnLysSerLeuGlnGlu 1504
Qy 1850 ----------GCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAA 1897
Db 1505 GluLeuSerLeuAlaArgGlyThrIleGluArgLeuThrLysSerLeuAlaAspValGlu 1524
Qy 1898 AATTCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCC 1957
Db 1525 SerGlnValSerAlaGlnAsn----------LysGluLysAspThrValLeuGly 1539
Qy 1958 AAA---GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCT 2014
Db 1540 ArgLeuAlaLeuLeuGlnGluGluArgAspLysLeuIleThrGluMetAspArgSerLeu 1559
Qy 2015 GTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCC 2074
Db 1560 LeuGluAsnGlnSerLeuSerSerSerCys----------GluSer 1571
Qy 2075 TTAGAATTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTG 2134
Db 1572 LeuLysLeu----------AlaLeuGluGlyLeuThrGluAspLysGluLysLeuVal 1587
Qy 2135 AAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGA 2194
Db 1588 LysGluIleGluSerLeuLysSerSerLysIleAlaGluSerThrGluTrpGlnGluLys 1607
Qy 2194 ---------- 2194
Db 1608 HisLysGluLeuGlnLysGluTyrGluIleLeuLeuGlnSerTyrGluAsnValSerAsn 1627
Qy 2195 ----------GAAACACTCAAAGCAGAGTCTCCTGATAATGAT 2227
Db 1628 GluAlaGluArgIleGlnHisValValGluAlaValArgGlnGluLysGlnGluLeuTyr 1647
Qy 2228 GGTCTTCTGAAGCCTACCTGTGGAAGGAAA---------- 2257
Db 1648 GlyLysLeuArgSerThrGluAlaAsnLysLysGluThrGluLysGlnLeuGlnGluAla 1667
Qy 2258 ----------GTTTCTCTTCCAAAT 2272
Db 1668 GluGlnGluMetGluGluMetLysGluLysMetArgLysPheAlaLysSerLysGlnGln 1687
Qy 2273 AAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAA 2332
Db 1688 LysIleLeuGluLeuGluGluGluAsnAspArgLeuArgAlaGluValHisProAlaGly 1707
Qy 2333 TCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTC 2392
Db 1708 AspThrAlaLysGlu----------CysMetGluThrLeu 1717
Qy 2393 TTACAGAATGATGTGTGTTTACCCAAGGCTACACAT------CAAAAAGAATTCGATACC 2446
Db 1718 LeuSerSerAsnAlaSerMetLysGluGluLeuGluArgValLysMetGluTyrGluThr 1737
Qy 2447 TTAAGTGGAAAATTAGAAGAG---TCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGT 2503
Db 1738 LeuSerLysLysPheGlnSerLeuMetSerGluLysAspSerLeuSerGluGluValGln 1757
Qy 2504 GGAATGAAAATTTCTCTTCCAAAT----------AAAGCCTTAGAA 2539
Db 1758 AspLeuLysHisGlnIleGluAspAsnValSerLysGlnAlaAsnLeuGluAlaThrGlu 1777
Qy 2540 TTGAAGGACAGAGAAACATTCAAAGCAGAGGAT----------GTGAGT 2578
Db 1778 LysHisAspAsnGlnThrAsnValThrGluGluGlyThrGlnSerIleProGlyGluThr 1797
Qy 2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACT---ACTGAAAATTCACAGTCT 2635
Db 1798 GluGluGlnAspSerLeuSerMetSerThrArgProThrCysSerGluSerValProSer 1817
Qy 2636 ACAAAA----------GTTGAGGAAGACTTTAATCTTACTACCAAG---GAGGGA 2677
Db 1818 AlaLysSerAlaAsnProAlaValSerLysAspPheSerSerHisAspGluIleAsnAsn 1837
Qy 2678 GCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCA 2737
Db 1838 TyrLeuGlnGlnIleAspGlnLeuLysGluArgIleAlaGlyLeuGluGlu---------- 1854
Qy 2738 CAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAA--- 2794
Db 1855 GluLysGlnLysAsnLysGluPheSerGlnThrLeuGluAsnGluLysAsnThrLeuLeu 1874
Qy 2795 ---------TCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAG 2845
Db 1875 SerGlnIleSerThrLysAspGlyGluLeuLysMetLeuGlnGluGluValThrLysMet 1894
Qy 2846 TGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAG 2905
Db 1895 AsnLeuLeuAsnGlnGlnIleGlnGluGluLeuSerArgValThr---LysLeuLysGlu 1913
Qy 2906 TCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAAT 2965
Db 1914 ThrAlaGlu----------GluGluLysAspAspLeuGluGluArgLeuMetAsn 1928
Qy 2966 AAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAA 3025
Db 1929 Gln----------LeuAlaGlu 1932
Qy 3026 ATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAA 3085
Db 1933 LeuAsnGlySerIle----------GlyAsnTyrCysGln 1942
Qy 3086 CAAATTACA---------GCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAA 3136
Db 1943 AspValThrAspAlaGlnIleLysAsnGluLeuLeuGluSerGluMetLysAsnLeuLys 1962

```
Qy       3137 AAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGG 3196
Db       1963 LysCysValSerGluLeuGluGluGluLysGlnGlnLeuValLysGluLysThrLysVal 1982

Qy       3197 GAACAAGAGCTCTGCAGTGTGAGATTG--------------------------CCTTTA 3229
Db       1983 GluSerGluIleArgLysGluTyrLeuGluLysIleGlnGlyAlaGlnLysGluProGly 2002

Qy       3230 AATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAA 3289
Db       2003 AsnLysSerHisAlaLysGluLeuGln---GluLeuLeuLysGluLysGlnGlnGluVal 2021

Qy       3290 GAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAA--------------CTTGAACAG 3334
Db       2022 LysGlnLeuGlnLysAspCys---IleArgTyrGlnGluLysIleSerAlaLeuGluArg 2040

Qy       3335 ACTCTCAGAATACAAGATATA--------------------GAATTGAAAAGTGTAACA 3373
Db       2041 ThrValLysAlaLeuGluPheValGlnThrGluSerGlnLysAspLeuGluIleThrLys 2060

Qy       3374 AGTAATTTGAATCAG---GTTTCTCACACTCATGAAAGTGAAAATGATCTC--------- 3421
Db       2061 GluAsnLeuAlaGlnAlaValGluHisArgLysLysAlaGlnAlaGluLeuAlaSerPhe 2080

Qy       3422 ----------------------------------------TTTCATGAAAATTGCATG 3439
Db       2081 LysValLeuLeuAspAspThrGlnSerGluAlaAlaArgValLeuAlaAspAsnLeuLys 2100

Qy       3440 TTGAAAAAGGAAATT-----------------------GCCATGCTAAAACTGGAAGTA 3475
Db       2101 LeuLysLysGluLeuGlnSerAsnLysGluSerValLysSerGlnMetLysGlnLysAsp 2120

Qy       3476 GCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATT 3535
Db       2121 GluAspLeuGluArgArgLeuGluGlnAlaGluGluLysHisLeuLysGluLysLysAsn 2140

Qy       3536 TTACAAGAAAAG----------------------------------------AATGCT 3553
Db       2141 MetGlnGluLysLeuAspAlaLeuArgArgGluLysValHisLeuGluGluThrIleGly 2160

Qy       3554 GAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTAT 3613
Db       2161 GluIleGlnValThrLeuAsnLysLysAspLysGluVal------------GlnGlnLeu 2176

Qy       3614 AGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTG 3667
Db       2177 GlnGluAsnLeuAspSerThrValThrGlnLeuAlaAlaPheThrLysSerMet 2194
```

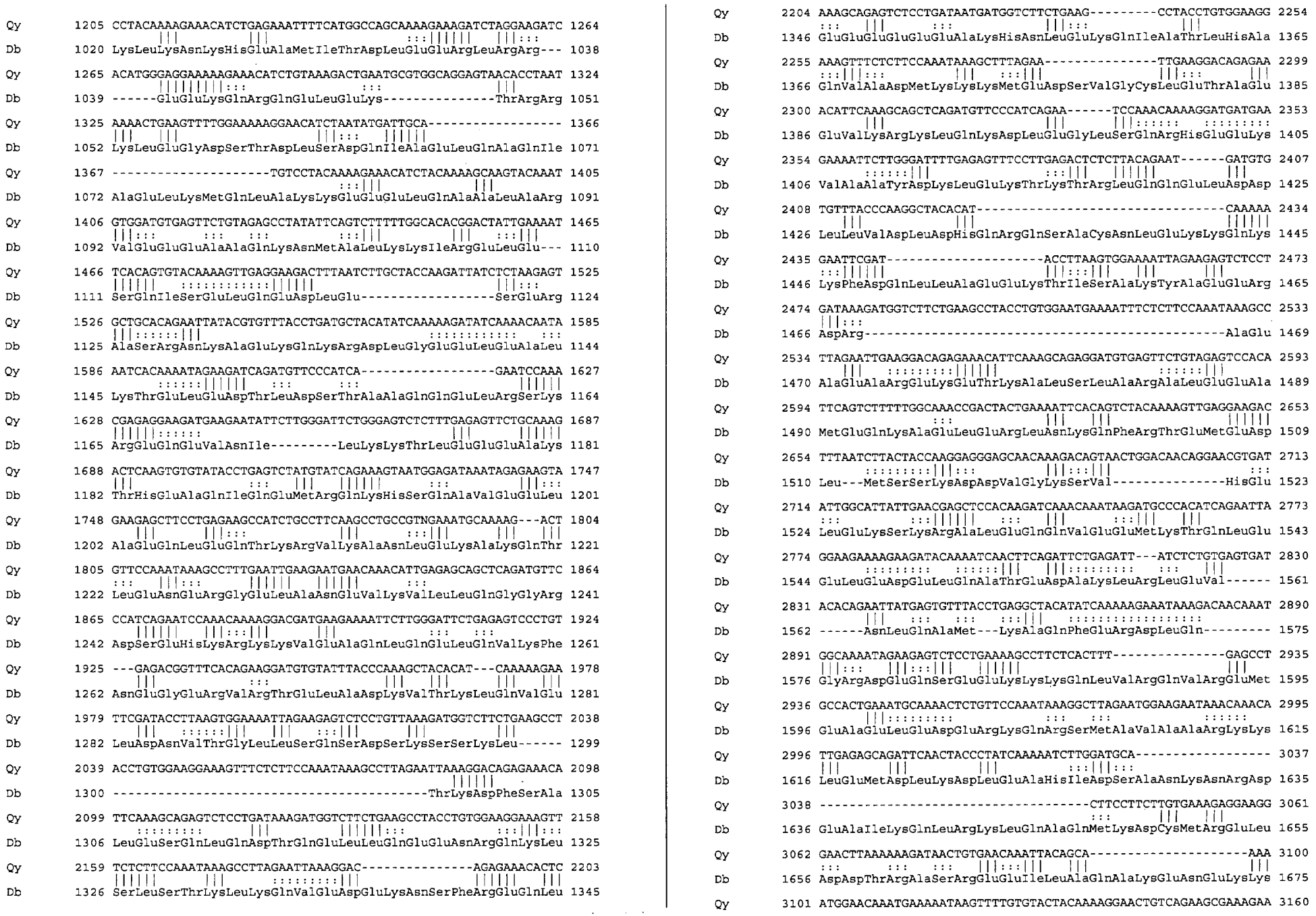
RESULT 13
A61231
myosin heavy chain nonmuscle form A - human
N;Alternate names: cellular myosin heavy chain; myosin type 9; NMMHC-A
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Homo sapiens (man)
C;Date: 12-May-1994 #sequence_revision 14-Jul-1994 #text_change 19-Apr-2002
C;Accession: A61231; A34876; I52562; I61692
R;Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Gdula, D.; Adelstein,
Circ. Res. 69, 530-539, 1991
A;Title: Human nonmuscle myosin heavy chains are encoded by two genes located on differe
A;Reference number: A61231; MUID:91316803; PMID:1860190
A;Accession: A61231
A;Molecule type: mRNA
A;Residues: 1-715 <SIM>
A;Cross-references: GB:M69180; NID:g189029; PIDN:AAA61765.1; PID:g189030
R;Saez, C.G.; Myers, J.C.; Shows, T.B.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 87, 1164-1168, 1990
A;Title: Human nonmuscle myosin heavy chain mRNA: generation of diversity through altern
A;Reference number: A34876; MUID:90138958; PMID:1967836
A;Accession: A34876
A;Molecule type: mRNA
A;Residues: 715-1961 <SAE>
A;Cross-references: GB:M31013; NID:g189035; PIDN:AAA36349.1; PID:g189036
R;Toothaker, L.E.; Gonzalez, D.A.; Tung, N.; Lemons, R.S.; Le Beau, M.M.; Arnaout, M.A.;
Blood 78, 1826-1833, 1991
A;Title: Cellular myosin heavy chain in human leukocytes: isolation of 5' cDNA clones, cl
A;Reference number: I52562; MUID:92003925; PMID:1912569
A;Accession: I52562
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-52,'EAI',56-659,'T',661-868,'T',870-930,'C',932-1239,'KG',1242-1337 <RES>
A;Cross-references: GB:M81105; NID:g188988; PIDN:AAA59888.1; PID:g553596
R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994
A;Title: Identification and overlapping expression of multiple unconventional myosin gene
A;Reference number: A55758; MUID:94294418; PMID:8022818
A;Accession: I61692
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 182-218 <BEM>
A;Cross-references: GB:L29141; NID:g457249; PIDN:AAA20904.1; PID:g531134
C;Genetics:
A;Gene: GDB:MYH9
A;Cross-references: GDB:120216; OMIM:160775
A;Map position: 22q12.3-22q13.1
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide
F;84-764/Domain: myosin motor domain homology <MMOT>
F;174-181/Region: nucleotide-binding motif A (P-loop)
F;552-565/Region: actin binding #status predicted
F;626-640/Region: actin binding #status predicted
F;837-1938/Domain: coiled coil #status predicted <COI>
F;837-1277/Domain: S2 #status predicted <DS2>
F;1278-1961/Domain: light meromyosin #status predicted <LMM>
F;1939-1961/Domain: carboxyl-terminal <CBT>
F;125/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F;180/Binding site: ATP (Lys) #status predicted
F;694,704/Active site: Cys #status predicted

Alignment Scores:

Pred. No.:	9.41e-10	Length:	1961
Score:	310.00	Matches:	223
Percent Similarity:	38.39%	Conservative:	192
Best Local Similarity:	20.63%	Mismatches:	444
Query Match:	4.81%	Indels:	222
DB:	1	Gaps:	43

US-09-602-362E-26 (1-3673) x A61231 (1-1961)

```
Qy        818 ATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTT 877
Db        895 LeuCysAlaGluAlaGluGluLeuArgAlaArgLeuThrAlaLysLysGlnGluLeu--- 913

Qy        878 GCTGAAGACATACATGGAATAACTGCA-----------------GAACGTTATGCTGCT 919
Db        914 GluGluIleCysHisAspLeuGluAlaArgValGluGluGluGluArgTyr------ 931

Qy        920 GCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAA 979
Db        932 ------------------------GlnHisLeuGlnAlaGluLysLysLysMetGlnGln 943

Qy        980 AATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCC 1039
Db        944 AsnIleGlnGluLeuGluGluGln---------LeuGluGluGluGluSerAlaArgGln 960

Qy       1040 TTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCT 1099
Db        961 LysLeuGlnLeuGluLysValThrThrGluAlaLysLeuLysLysLeuGluGluGluGln 980

Qy       1100 GCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGA 1159
Db        981 Ile---IleLeuGluAspGlnAsnCysLysLeuAlaLysGluLysLysLeuGluAsp 999

Qy       1160 AAGTTTGAACAGTCAACA--------------GAAGAAACACCTAGGAAAATTTTGAGG 1204
Db       1000 ArgIleAlaGluPheThrThrAsnLeuThrGluGluGluGluLysSerLysSerLeuAla 1019
```



```
Db       1676 LeuLysSerMetGluAlaGluMetIleGlnLeuGlnGluGluLeuAlaAlaAlaGluArg 1695

Qy       3161 ATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTG--- 3217

Db       1696 AlaLysArgGlnAlaGlnGlnGluArgAspGluLeuAlaAspGluIleAlaAsnSerSer 1715

Qy       3218 ---------AGATTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTA 3268

Db       1716 GlyLysGlyAlaLeuAlaLeu---------GluGluLysArgArg-----------Leu 1728

Qy       3269 AAAGAAAAAATT---AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAA 3325

Db       1729 GluAlaArgIleAlaGlnLeuGluGluGluLeuGluGluGluGlnGlyAsnThrGluLeu 1748

Qy       3326 CTTGAACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAAT 3385

Db       1749 IleAsnAspArgLeuLysLysAlaAsnLeuGlnIleAspGlnIleAsnAlaAspLeuAsn 1768

Qy       3386 ---CAGGTTTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTG 3442

Db       1769 LeuGluArgGlyHisAlaGlnLysAsnGluAsn-----------AlaArgGlnGlnLeu 1784

Qy       3443 AAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTG 3502

Db       1785 GluArgGlnAsnLysGluLeuLysValLys---------LeuGlnGluMetGluGlyThr 1801

Qy       3503 AAGGAAAATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAA 3562

Db       1802 ValLysSerLysTyrLysAlaSerIleThrAlaLeuGluAlaLysIleAlaGlnLeuGlu 1821

Qy       3563 ATGACCCTA---------------------------------------------------- 3571

Db       1822 GluGlnLeuAspAsnGluThrLysGluArgGlnAlaAlaCysLysGlnValArgArgThr 1841

Qy       3572 ------AAACTGAAACAGAAAACAGTA--------------ACAAAAAGGGCATCTCAG 3610

Db       1842 GluLysLysLeuLysAspValLeuLeuGlnValAspAspGluArgArgAsnAlaGluGln 1861

Qy       3611 TATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAG 3670

Db       1862 TyrLysAspGlnAlaAspLysAlaSerThrArgLeuLysGlnLeuLysArgGlnLeuGlu 1881

Qy       3671 GAA 3673

Db       1882 Glu 1882
```

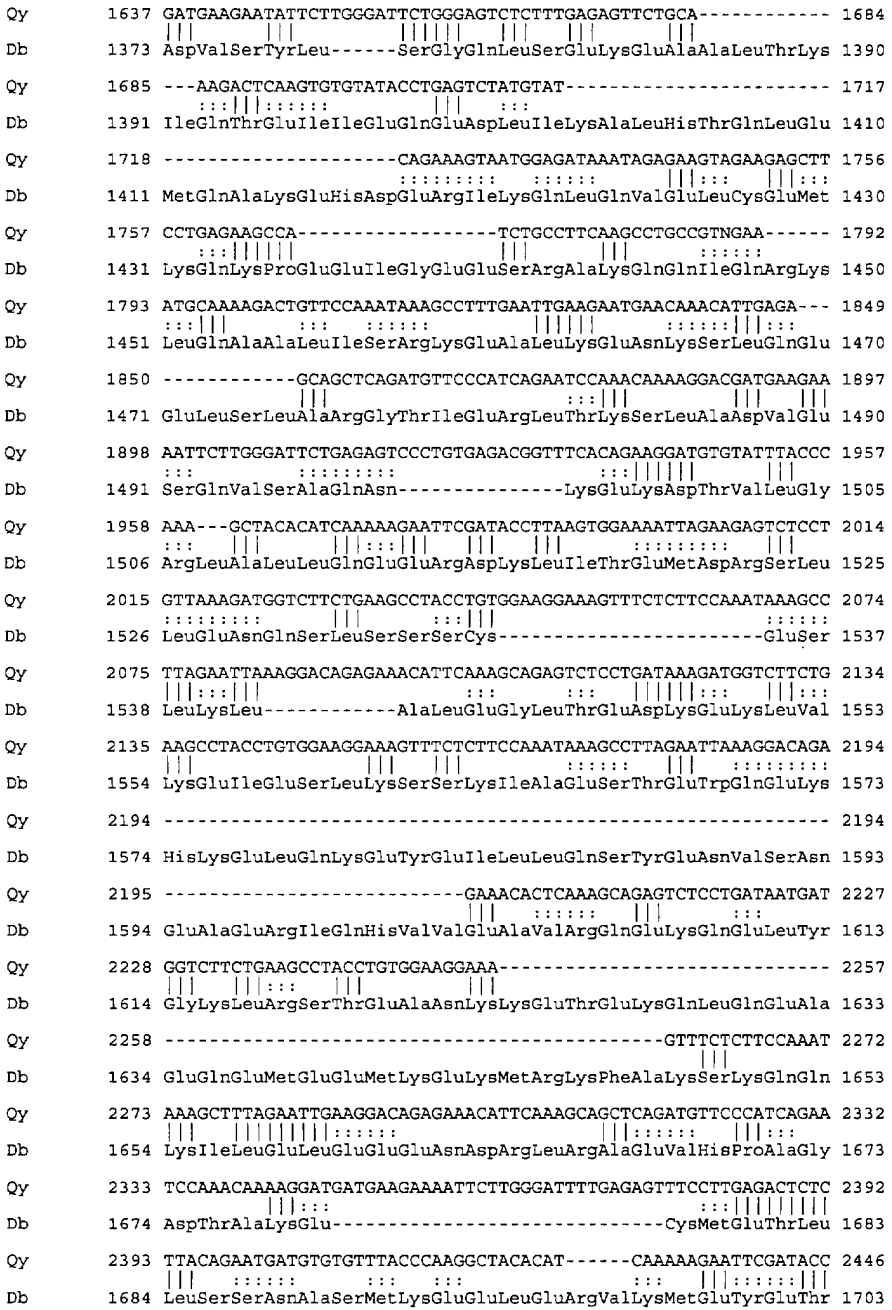
RESULT 14
I52300
giantin - human
N;Alternate names: gcp372
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 20-Aug-1999
C;Accession: I52300
R;Sohda, M.; Misumi, Y.; Fujiwara, T.; Nishioka, M.; Ikehara, Y.
Biochem. Biophys. Res. Commun. 205, 1399-1408, 1994
A;Title: Molecular cloning and sequence analysis of a human 372-kDA protein localized in
A;Reference number: I52300; MUID:95100974; PMID:7802676
A;Accession: I52300
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3225 <RES>
A;Cross-references: GB:D25542; NID:g662389; PIDN:BAA05025.1; PID:g808869
C;Superfamily: giantin

Alignment Scores:

Pred. No.:	9.95e-10	Length:	3225
Score:	310.00	Matches:	235
Percent Similarity:	36.62%	Conservative:	211
Best Local Similarity:	19.29%	Mismatches:	426
Query Match:	4.81%	Indels:	346
DB:	2	Gaps:	48

US-09-602-362E-26 (1-3673) x I52300 (1-3225)

```
Qy        680 GTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAA 739

Db       1067 LeuGlnAspLysThrAsnGlnIleAspLeuLeuGlnAlaGluIleSerGluAsnGlnAla 1086

Qy        740 ACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAG--------- 790

Db       1087 IleIleGlnLysLeuIleThrSerAsnThrAspAlaSerAspGlyAspSerValAlaLeu 1106

Qy        791 TCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATG 850

Db       1107 ValLysGluThrValValIleSerProProCysThrGlySerSerGluHisTrpLysPro 1126

Qy        851 CTTCTTCAGCAAAATGTT---------GACGTCTTTGCTGAAGACATACATGGAATAACT 901

Db       1127 GluLeuGluGluLysIleLeuAlaLeuGluLysGluLysGluGlnLeuGlnLysLysLeu 1146

Qy        902 GCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACAT 961

Db       1147 GlnGluAlaLeuThrSerArgLysAlaIleLeuLysLysAlaGlnGluLysGluArgHis 1166

Qy        962 ATACGA---------AAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCT 1012

Db       1167 LeuArgGluGluLeuLysGlnGlnLysAspAspTyrAsnArgLeuGlnGluGlnPhe--- 1185

Qy       1013 ACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGC 1072

Db       1186 ------------AspGluGlnSerLysGluAsnGluAsnIleGlyAspGlnLeuArgGln 1201

Qy       1073 TTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATT 1132

Db       1202 LeuGlnIleGlnValArgGluSerIleAspGlyLysLeuProSerThrAspGlnGlnGlu 1221

Qy       1133 CAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGG 1192

Db       1222 SerCys------SerSerThrProGlyLeuGluGluProLeuPheLysAlaThrGluGln 1239

Qy       1193 AAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGA 1252

Db       1240 HisHisThrGlnProValLeuGluSerAsnLeuCysProAspTrpProSerHisSerGlu 1259

Qy       1253 TCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGA 1312

Db       1260 Asp---------------------------------------------------------- 1260

Qy       1313 GTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCT---AATATGATTGCATGT 1369

Db       1261 ---------------AlaSerAlaLeuGlnGlyGlyThrSerValAlaGlnIleLysAla 1275

Qy       1370 CCTACAAAAGAA---ACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAG 1426

Db       1276 GlnLeuLysGluIleGluAlaGluLysValGluLeuGluLeuLysValSerSerThr--- 1294

Qy       1427 CCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAG 1486

Db       1295 ----------------------------------ThrSerGluLeuThrLysLysSer 1302

Qy       1487 GAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGT 1546

Db       1303 GluGluValPheGlnLeuGlnGluGlnIleAsnLysGlnGlyLeuGlu------------ 1318

Qy       1547 TTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA------ 1600

Db       1319 -----------------IleGluSerLeuLysThrValSerHisGluAlaGluValHis 1332

Qy       1600 ---------------------------------------------------------- 1600

Db       1333 AlaGluSerLeuGlnGlnLysLeuGluSerSerGlnLeuGlnIleAlaGlyLeuGluHis 1352

Qy       1601 ----------------------GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAA 1636

Db       1353 LeuArgGluLeuGlnProLysLeuAspGluLeuGlnLysLeuIleSerLysLysGluGlu 1372
```


Qy 1637 GATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGCA---------- 1684
Db 1373 AspValSerTyrLeu------SerGlyGlnLeuSerGluLysGluAlaAlaLeuThrLys 1390
Qy 1685 ---AAGACTCAAGTGTGTATACCTGAGTCTATGTAT---------- 1717
Db 1391 IleGlnThrGluIleIleGluGlnGluAspLeuIleLysAlaLeuHisThrGlnLeuGlu 1410
Qy 1718 ----------CAGAAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTT 1756
Db 1411 MetGlnAlaLysGluHisAspGluArgIleLysGlnLeuGlnValGluLeuCysGluMet 1430
Qy 1757 CCTGAGAAGCCA----------TCTGCCTTCAAGCCTGCCGTNGAA------ 1792
Db 1431 LysGlnLysProGluGluIleGlyGluGluSerArgAlaLysGlnGlnIleGlnArgLys 1450
Qy 1793 ATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGA--- 1849
Db 1451 LeuGlnAlaAlaLeuIleSerArgLysGluAlaLeuLysGluAsnLysSerLeuGlnGlu 1470
Qy 1850 ----------GCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAA 1897
Db 1471 GluLeuSerLeuAlaArgGlyThrIleGluArgLeuThrLysSerLeuAlaAspValGlu 1490
Qy 1898 AATTCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCC 1957
Db 1491 SerGlnValSerAlaGlnAsn----------LysGluLysAspThrValLeuGly 1505
Qy 1958 AAA---GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCT 2014
Db 1506 ArgLeuAlaLeuLeuGlnGluGluArgAspLysLeuIleThrGluMetAspArgSerLeu 1525
Qy 2015 GTTAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCC 2074
Db 1526 LeuGluAsnGlnSerLeuSerSerSerCys----------GluSer 1537
Qy 2075 TTAGAATTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTG 2134
Db 1538 LeuLysLeu----------AlaLeuGluGlyLeuThrGluAspLysGluLysLeuVal 1553
Qy 2135 AAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGA 2194
Db 1554 LysGluIleGluSerLeuLysSerSerLysIleAlaGluSerThrGluTrpGlnGluLys 1573
Qy 2194 ---------- 2194
Db 1574 HisLysGluLeuGlnLysGluTyrGluIleLeuLeuGlnSerTyrGluAsnValSerAsn 1593
Qy 2195 ----------GAAACACTCAAAGCAGAGTCTCCTGATAATGAT 2227
Db 1594 GluAlaGluArgIleGlnHisValValGluAlaValArgGlnGluLysGlnGluLeuTyr 1613
Qy 2228 GGTCTTCTGAAGCCTACCTGTGGAAGGAAA---------- 2257
Db 1614 GlyLysLeuArgSerThrGluAlaAsnLysLysGluThrGluLysGlnLeuGlnGluAla 1633
Qy 2258 ----------GTTTCTCTTCCAAAT 2272
Db 1634 GluGlnGluMetGluGluMetLysGluLysMetArgLysPheAlaLysSerLysGlnGln 1653
Qy 2273 AAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAA 2332
Db 1654 LysIleLeuGluLeuGluGluGluAsnAspArgLeuArgAlaGluValHisProAlaGly 1673
Qy 2333 TCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTC 2392
Db 1674 AspThrAlaLysGlu----------CysMetGluThrLeu 1683
Qy 2393 TTACAGAATGATGTGTGTTTACCCAAGGCTACACAT------CAAAAAGAATTCGATACC 2446
Db 1684 LeuSerSerAsnAlaSerMetLysGluGluLeuGluArgValLysMetGluTyrGluThr 1703
Qy 2447 TTAAGTGGAAAATTAGAAGAG---TCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGT 2503
Db 1704 LeuSerLysLysPheGlnSerLeuMetSerGluLysAspSerLeuSerGluGluValGln 1723
Qy 2504 GGAATGAAAATTTCTCTTCCAAAT----------AAAGCCTTAGAA 2539
Db 1724 AspLeuLysHisGlnIleGluGlyAsnValSerLysGlnAlaAsnLeuGluAlaThrGlu 1743
Qy 2540 TTGAAGGACAGAGAAACATTCAAAGCAGAGGAT----------GTGAGT 2578
Db 1744 LysHisAspAsnGlnThrAsnValThrGluGluGlyThrGlnSerIleProGlyGluThr 1763
Qy 2579 TCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACT---ACTGAAAATTCACAGTCT 2635
Db 1764 GluGluGlnAspSerLeuSerMetSerThrArgProThrCysSerGluSerValProSer 1783
Qy 2636 ACAAAA----------GTTGAGGAAGACTTTAATCTTACTACCAAG---GAGGGA 2677
Db 1784 AlaLysSerAlaAsnProAlaValSerLysAspPheSerSerHisAspGluIleAsnAsn 1803
Qy 2678 GCAACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCA 2737
Db 1804 TyrLeuGlnGlnIleAspGlnLeuLysGluArgIleAlaGlyLeuGluGlu---------- 1820
Qy 2738 CAAGATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAA--- 2794
Db 1821 GluLysGlnLysAsnLysGluPheSerGlnThrLeuGluAsnGluLysAsnThrLeuLeu 1840
Qy 2795 ---------TCAACTTCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAG 2845
Db 1841 SerGlnIleSerThrLysAspGlyGluLeuLysMetLeuGlnGluGluValThrLysMet 1860
Qy 2846 TGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAG 2905
Db 1861 AsnLeuLeuAsnGlnGlnIleGlnGluGluLeuSerArgValThr---LysLeuLysGlu 1879
Qy 2906 TCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAAT 2965
Db 1880 ThrAlaGlu----------GluGluLysAspAspLeuGluGluArgLeuMetAsn 1894
Qy 2966 AAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAA 3025
Db 1895 Gln----------LeuAlaGlu 1898
Qy 3026 ATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAA 3085
Db 1899 LeuAsnGlySerIle----------GlyAsnTyrCysGln 1908
Qy 3086 CAAATTACA---------GCAAAAATGGAACAAATGAAAAATAAGTTTTGTGTACTACAA 3136
Db 1909 AspValThrAspAlaGlnIleLysAsnGluLeuLeuGluSerGluMetLysAsnLeuLys 1928
Qy 3137 AAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGG 3196
Db 1929 LysCysValSerGluLeuGluGluGluLysGlnGlnLeuValLysGluLysThrLysVal 1948
Qy 3197 GAACAAGAGCTCTGCAGTGTGAGATTG----------CCTTTA 3229
Db 1949 GluSerGluIleArgLysGluTyrLeuGluLysIleGlnGlyAlaGlnLysGluProGly 1968
Qy 3230 AATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAA 3289
Db 1969 AsnLysSerHisAlaLysGluLeuGln---GluLeuLeuLysGluLysGlnGlnGluVal 1987
Qy 3290 GAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAA----------CTTGAACAG 3334
Db 1988 LysGlnLeuGlnLysAspCys---IleArgTyrGlnGluLysIleSerAlaLeuGluArg 2006
Qy 3335 ACTCTCAGAATACAAGATATA----------GAATTGAAAAGTGTAACA 3373
Db 2007 ThrValLysAlaLeuGluPheValGlnThrGluSerGlnLysAspLeuGluIleThrLys 2026
Qy 3374 AGTAATTTGAATCAG---GTTTCTCACACTCATGAAAGTGAAAATGATCTC--------- 3421

```
                ||||||   |||   |||   |||      :::::::::   :::|||
Db       2027 GluAsnLeuAlaGlnAlaValGluHisArgLysLysAlaGlnAlaGluLeuAlaSerPhe 2046

Qy       3422 ----------------------------------------TTTCATGAAAATTGCATG 3439
                                                       :::|||
Db       2047 LysValLeuLeuAspAspThrGlnSerGluAlaAlaArgValLeuAlaAspAsnLeuLys 2066

Qy       3440 TTGAAAAAGGAAATT----------------------GCCATGCTAAAACTGGAAGTA 3475
              ||||||||||||:::                 :::   :::|||   :::
Db       2067 LeuLysLysGluLeuGlnSerAsnLysGluSerValLysSerGlnMetLysGlnLysAsp 2086

Qy       3476 GCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATT 3535
              |||:::   :::  :::     |||   |||:::   ::::::   |||
Db       2087 GluAspLeuGluArgArgLeuGluGlnAlaGluGluLysHisLeuLysGluLysLysAsn 2106

Qy       3536 TTACAAGAAAAG----------------------------------------AATGCT 3553
              :::|||||||||
Db       2107 MetGlnGluLysLeuAspAlaLeuArgArgGluLysValHisLeuGluGluThrIleGly 2126

Qy       3554 GAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTAT 3613
              |||:::|||:::||||||      |||  |||  |||                |||
Db       2127 GluIleGlnValThrLeuAsnLysLysAspLysGluVal------------GlnGlnLeu 2142

Qy       3614 AGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTG 3667
              :::|||   |||                :::             |||      :::
Db       2143 GlnGluAsnLeuAspSerThrValThrGlnLeuAlaAlaPheThrLysSerMet 2160

RESULT 15
JC5837
364K Golgi complex-associated protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
C;Accession: JC5837
R;Toki, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.
Cell Struct. Funct. 22, 565-577, 1997
A;Title: Identification and characterization of rat 364-kDa Golgi-associated protein rec
A;Reference number: JC5837; MUID:98093490; PMID:9431462
A;Accession: JC5837
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3187 <TOK>
A;Cross-references: DDBJ:D25543; NID:g516825; PIDN:BAA05026.1; PID:g516826
C;Comment: This protein plays a role in the formation and maintenance of the characteris
C;Superfamily: giantin
F;49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predict
F;3165-3187/Domain: membrane anchor #status predicted <MAD>

Alignment Scores:
Pred. No.:              2.59e-09        Length:         3187
Score:                  302.50          Matches:        253
Percent Similarity:     37.56%          Conservative:   197
Best Local Similarity:  21.12%          Mismatches:     451
Query Match:            4.69%           Indels:         297
DB:                     2               Gaps:           54

US-09-602-362E-26 (1-3673) x JC5837 (1-3187)

Qy        641 ATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAG--------- 691
              :::     :::|||          :::   :::   |||||||||:::
Db       1453 LeuThrLysSerLeuAlaAspValGluSerGlnValSerValGlnAsnGlnGluLysAsp 1472

Qy        692 GCTAGCCTCACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTT 751
              |||   |||       |||   |||      ::::::|||   |||      |||||
Db       1473 AlaLeuLeuGlyLysLeuAlaLeuLeuGlnGluGluArgAspLysLeuIleValGlu--- 1491

Qy        752 TTACTAACAAAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATG 811
              :::   |||:::          |||                |||   :::|||
Db       1492 ---MetAspLysSerLeuLeuGluAsnGlnSerLeuGlyGlySerCysGluSerLeuLys 1510

Qy        812 CTTGCCATATGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGAC 871
              ||||||:::      |||                |||   ::::::   :::
Db       1511 LeuAlaLeu------GlyGly--------------------LeuThrGluAspLysGlu 1521

Qy        872 GTCTTTGCTGAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTT 931
                          :::::::::      :::      :::   |||   :::
Db       1522 LysLeuMetLysGluLeuGluSerValArgCysSerLysIleAlaGluSerThrGluTrp 1541

Qy        932 AATTACATTCATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAAT 991
                   |||:::::::|||   :::          |||   ::::::   :::|||
Db       1542 GlnGluLysHisLysGluLeuGlnLysGluTyrGluValLeuLeuGlnSerTyrGluAsn 1561

Qy        992 ACCAATCCAGAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGA 1051
              :::   |||                                          |||||||||
Db       1562 ValSerAsnGlu----------------------------------------AlaGluArg 1568

Qy       1052 ACACCTGACACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTG 1111
                        |||||:::   ::::::      :::      |||:::|||
Db       1569 IleGlnHisValValGluSerValArgGlnGluLysGlnGluValTyrAlaLysLeuArg 1588

Qy       1112 GAGGGAACGTCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAG 1171
                   |||  |||  :::          :::                  ||||||
Db       1589 SerAlaGluSerAspLysArgGluArgGluLysGlnLeuGlnAspAla------GluGln 1606

Qy       1172 TCAACAGAAGAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTT 1231
                 ||||||        |||:::   :::      |||          :::|||
Db       1607 GluMetGluGluMetLysGluLysMetArgLysPheAlaLysSerLysGlnGlnLysIle 1626

Qy       1231 ------------------------------------------------------------ 1231

Db       1627 LeuGluLeuGluGluGluAsnAspArgLeuArgAlaGluAlaGlnProValGlyGlyAla 1646

Qy       1232 ----------------------------TCATGGCCAGCAAAAGAAAGATCTAGGAAG 1261
                                         :::        ||||||        :::
Db       1647 AsnGluSerMetGluAlaLeuLeuSerSerAsnAlaSerLeuLysGluGluLeuGluArg 1666

Qy       1262 ATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCT 1321
              ||||||   |||  |||     |||  :::  |||   :::
Db       1667 IleThrLeuGluTyrLysThrLeuSerLysGluPheGluAlaLeuMetAla--------- 1683

Qy       1322 AATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATG---------------------- 1360
                 :::         |||  :::  |||  |||:::
Db       1684 ---GluLysAsnThrLeuSerGluGluThrArgAsnLeuLysLeuGlnValGluAlaGln 1702

Qy       1361 ---ATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGT 1417
                 :::         :::  ||||||:::  |||:::          ||||||
Db       1703 GluLeuLysGlnAlaSerLeuGluThrThrGluLysSerAspGluProLysAspVal--- 1721

Qy       1418 TCTGTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAAT---------TCA 1468
                 :::|||  :::                :::::   |||          |||
Db       1722 ---IleGluGluValThrGluAlaValValGlyLysSerGlnGluGlnAspSerLeuSer 1740

Qy       1469 CAGTGTACAAAAGTTGAGGAA---GACTTTAATCTTGCTACCAAGATTATCTCTAAGAGT 1525
                 :::      |||:::|||:::   :::     |||
Db       1741 GluAsnAlaLysLeuGluAspAlaGluAlaThrLeuLeuAlaAsnSerAlaLysProGly 1760

Qy       1526 GCTGCACAGAATTATACGTGTTTACCTGATGCT---ACATATCAAAAAGATATCAAAACA 1582
                 ::::::   ::::::         |||          |||   :::   :::
Db       1761 ValSerGluThrPheSerSerHisAspAspIleAsnAsnTyrLeuGlnGlnLeuAspGln 1780

Qy       1583 ATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAA 1642
                 :::      ::::|||   :::                |||   ::::::::::::|||
Db       1781 LeuLysGlyArgIleAlaGlu------------LeuGluMetGluLysGlnLysAspArg 1796

Qy       1643 GAATATTCTTGG--------------GATTCTGGGAGTCTCTTTGAGAGTTCTGCAAAG 1687
              |||   |||                     :::      |||   :::   |||||||||
Db       1797 GluLeuSerGlnThrLeuGluAsnGluLysAsnAlaLeuLeuThrGlnIleSerAlaLys 1816

Qy       1688 ---ACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGAA 1744
                 :::::::::   :::  |||           |||:::      :::|||::::::
Db       1817 AspSerGluLeuLysLeuGluGluGluValAlaLysIleAsnMetLeuAsnGlnGln 1836

Qy       1745 GTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAA------------ 1792
                 ::::::|||      :::  :::      |||      |||
Db       1837 IleGlnGluGluLeuSerArgValThrLysLeuLysGluThrAlaGluGluGluLysAsp 1856

Qy       1793 ---ATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTG---------------------- 1828
```

```
Qy       3623 ------CTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAG 3670
                  |||      |||      ||||||      |||:::   :::::::|||
Db       2524 MetSerLeuAspAlaLeuGlnGluGluAsnGlnGlyLeuSerLysGluIleLys 2541
```

Search completed: July 15, 2004, 09:51:56
Job time : 233.685 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 15, 2004, 08:50:37 ; Search time 171.903 Seconds
(without alignments)
13356.780 Million cell updates/sec

Title: US-09-602-362E-26
Perfect score: 6448
Sequence: 1 caagagcttggcgatacaga.....tgacttctaaattgaaggaa 3673

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 2570690

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09602362/runat_15072004_093627_22071/app_query.fasta_1.10325
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09602362_@CGN_1_1_719_@runat_15072004_093627_22071
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5938	92.1	1225	14	US-10-177-293-332	Sequence 332, App
2	5173	80.2	1011	16	US-10-408-765A-1557	Sequence 1557, Ap
3	5173	80.2	1239	13	US-10-007-805-577	Sequence 577, App
4	5173	80.2	1239	14	US-10-076-622-577	Sequence 577, App
5	5173	80.2	1239	14	US-10-124-805-577	Sequence 577, App
6	3458	53.6	1341	12	US-10-058-270A-4	Sequence 4, Appli
7	3458	53.6	1341	13	US-10-007-805-565	Sequence 565, App
8	3458	53.6	1341	14	US-10-076-622-565	Sequence 565, App
9	3458	53.6	1341	14	US-10-177-293-334	Sequence 334, App
10	3458	53.6	1341	14	US-10-124-805-565	Sequence 565, App
11	3440	53.3	1349	13	US-10-007-805-573	Sequence 573, App
12	3440	53.3	1349	14	US-10-076-622-573	Sequence 573, App
13	3440	53.3	1349	14	US-10-124-805-573	Sequence 573, App
14	2374	36.8	1013	13	US-10-007-805-553	Sequence 553, App
15	2374	36.8	1013	14	US-10-076-622-553	Sequence 553, App
16	2374	36.8	1013	14	US-10-124-805-553	Sequence 553, App
17	2372.5	36.8	1095	13	US-10-007-805-493	Sequence 493, App
18	2372.5	36.8	1095	14	US-10-076-622-493	Sequence 493, App
19	2372.5	36.8	1095	14	US-10-124-805-493	Sequence 493, App
20	2367	36.7	1002	9	US-09-604-287A-475	Sequence 475, App
21	2367	36.7	1002	10	US-09-551-621-475	Sequence 475, App
22	2367	36.7	1002	13	US-10-007-805-475	Sequence 475, App
23	2367	36.7	1002	14	US-10-076-622-475	Sequence 475, App
24	2367	36.7	1002	14	US-10-124-805-475	Sequence 475, App
25	1776.5	27.6	661	13	US-10-007-805-552	Sequence 552, App
26	1776.5	27.6	661	14	US-10-076-622-552	Sequence 552, App
27	1776.5	27.6	661	14	US-10-124-805-552	Sequence 552, App
28	1775	27.5	743	13	US-10-007-805-494	Sequence 494, App
29	1775	27.5	743	14	US-10-076-622-494	Sequence 494, App
30	1775	27.5	743	14	US-10-124-805-494	Sequence 494, App
31	1769.5	27.4	650	9	US-09-825-301-25	Sequence 25, Appl
32	1769.5	27.4	650	9	US-09-604-287A-469	Sequence 469, App
33	1769.5	27.4	650	10	US-09-551-621-469	Sequence 469, App
34	1769.5	27.4	650	13	US-10-007-805-469	Sequence 469, App
35	1769.5	27.4	650	14	US-10-076-622-469	Sequence 469, App
36	1769.5	27.4	650	14	US-10-124-805-469	Sequence 469, App
37	1769.5	27.4	650	14	US-10-033-527-25	Sequence 25, Appl
38	1542	23.9	512	16	US-10-181-663-16	Sequence 16, Appl
39	1414.5	21.9	445	9	US-09-825-301-29	Sequence 29, Appl
40	1414.5	21.9	445	9	US-09-604-287A-473	Sequence 473, App
41	1414.5	21.9	445	10	US-09-551-621-473	Sequence 473, App
42	1414.5	21.9	445	13	US-10-007-805-473	Sequence 473, App
43	1414.5	21.9	445	14	US-10-076-622-473	Sequence 473, App
44	1414.5	21.9	445	14	US-10-124-805-473	Sequence 473, App
45	1414.5	21.9	445	14	US-10-033-527-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-10-177-293-332
; Sequence 332, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,

; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-332

Alignment Scores:

Pred. No.:	0	Length:	1225
Score:	5938.00	Matches:	1159
Percent Similarity:	99.57%	Conservative:	1
Best Local Similarity:	99.48%	Mismatches:	4
Query Match:	92.09%	Indels:	1
DB:	14	Gaps:	0

US-09-602-362E-26 (1-3673) x US-10-177-293-332 (1-1225)

```
Qy        181 ATGAAGAGGCTCTTAGCTGCCGCTGGCAAGGGCGTGCGGGGCCCGGAGCCCCCGAACCCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLysArgLeuLeuAlaAlaAlaGlyLysGlyValArgGlyProGluProProAsnPro 20

Qy        241 TTCAGCGAACGGGTCTACACTGAGAAGGACTACGGGACCATCTACTTCGGGGATCTAGGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 PheSerGluArgValTyrThrGluLysAspTyrGlyThrIleTyrPheGlyAspLeuGly 40

Qy        301 AAGATCCATACAGCTGCCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAGATGACAGTAGGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysIleHisThrAlaAlaSerArgGlyGlnValGlnLysLeuGluLysMetThrValGly 60

Qy        361 AAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTACACTGGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LysLysProValAsnLeuAsnLysArgAspMetLysLysArgThrAlaLeuHisTrpAla 80

Qy        421 TGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGCCNGCTTAAT 480
              |||||||||||||||||  |||||||||||||||||||||||||||||  ||||||
Db         81 CysValAsnGlyHisAlaGluValValThrPheLeuValAspArgLysCysGlnLeuAsn 100

Qy        481 GTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAGAGGGAAGCT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ValLeuAspGlyGluGlyArgThrProLeuMetLysAlaLeuGlnCysGluArgGluAla 120

Qy        541 TTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGTGTATGGCAA 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 -CysAlaAsnIleLeuIleAspAlaGlyAlaAspLeuAsnTyrValAspValTyrGlyAs 140

Qy        601 CACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAACACTGCTGTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 nThrAlaLeuHisTyrAlaValTyrSerGluAsnLeuLeuMetValAlaThrLeuLeuSe 160

Qy        661 CTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTTACTGGCCAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 rTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeuThrProLeuLeuLeuAlaIl 180

Qy        721 ACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAATGCAAACGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 eGlnLysArgSerLysGlnThrValGluPheLeuLeuThrLysAsnAlaAsnAlaAsnAl 200

Qy        781 ATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTCATCAGAGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 aPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIleCysGluGlySerSerGluIl 220

Qy        841 AGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACATGGAATAAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 eValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaGluAspIleHisGlyIleTh 240

Qy        901 TGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACTTTTGGAACA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 rAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIleHisGlnGlnLeuLeuGluHi 260

Qy        961 TATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATCTACAGGAAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 sIleArgLysLeuProLysAsnProGlnAsnThrAsnProGluGlyThrSerThrGlyTh 280

Qy       1021 ACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGACACGGCTGAAAGCTTGCTGGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 rProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGluSerLeuLeuGl 300

Qy       1081 AAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGCCAAAATTCAATGTCT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 uLysThrProAspGluAlaAlaArgLeuValGluGlyThrSerAlaLysIleGlnCysLe 320

Qy       1141 GGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAACACCTAGGAAAATTTT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 uGlyLysAlaThrSerGlyLysPheGluGlnSerThrGluGluThrProArgLysIleLe 340

Qy       1201 GAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGGAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 uArgProThrLysGluThrSerGluLysPheSerTrpProAlaLysGluArgSerArgLy 360

Qy       1261 GATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 sIleThrTrpGluGluLysGluThrSerValLysThrGluCysValAlaGlyValThrPr 380

Qy       1321 TAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 oAsnLysThrGluValLeuGluLysGlyThrSerAsnMetIleAlaCysProThrLysGl 400

Qy       1381 AACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 uThrSerThrLysAlaSerThrAsnValAspValSerSerValGluProIlePheSerLe 420

Qy       1441 TTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 uPheGlyThrArgThrIleGluAsnSerGlnCysThrLysValGluGluAspPheAsnLe 440

Qy       1501 TGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTAC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 uAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyrThrCysLeuProAspAlaTh 460

Qy       1561 ATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 rTyrGlnLysAspIleLysThrIleAsnHisLysIleGluAspGlnMetProSerGl 480

Qy       1621 ATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 uSerLysArgGluGluAspGluGluTyrSerTrpAspSerGlySerLeuPheGluSerSe 500

Qy       1681 TGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 rAlaLysThrGlnValCysIleProGluSerMetTyrGlnLysValMetGluIleAsnAr 520

Qy       1741 AGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db        520 gGluValGluGluLeuProGluLysProSerAlaPheLysProAlaIleGluMetGlnLy 540

Qy       1801 GACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGAT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 sThrValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAlaGlnMe 560
```

```
Qy       1861 GTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 tPheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspSerGluSerPr 580

Qy       1921 CTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATT 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 oCysGluThrValSerGlnLysAspValTyrLeuProLysAlaThrHisGlnLysGluPh 600

Qy       1981 CGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTAC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 eAspThrLeuSerGlyLysLeuGluGluSerProValLysAspGlyLeuLeuLysProTh 620

Qy       2041 CTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATT 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 rCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPh 640

Qy       2101 CAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 eLysAlaGluSerProAspLysAspGlyLeuLeuLysProThrCysGlyArgLysValSe 660

Qy       2161 TCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGA 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 rLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrLeuLysAlaGluSerProAs 680

Qy       2221 TAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 pAsnAspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLe 700

Qy       2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 uGluLeuLysAspArgGluThrPheLysAlaAlaGlnMetPheProSerGluSerLysGl 720

Qy       2341 AAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAA 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 nLysAspAspGluGluAsnSerTrpAspPheGluSerPheLeuGluThrLeuLeuGlnAs 740

Qy       2401 TGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 nAspValCysLeuProLysAlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLe 760

Qy       2461 AGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 uGluGluSerProAspLysAspGlyLeuLeuLysProThrCysGlyMetLysIleSerLe 780

Qy       2521 TCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 uProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaGluAspValSerSe 800

Qy       2581 TGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 rValGluSerThrPheSerLeuPheGlyLysProThrThrGluAsnSerGlnSerThrLy 820

Qy       2641 AGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 sValGluGluAspPheAsnLeuThrThrLysGluGlyAlaThrLysThrValThrGlyGl 840

Qy       2701 ACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCC 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 nGlnGluArgAspIleGlyIleIleGluArgAlaProGlnAspGlnThrAsnLysMetPr 860

Qy       2761 CACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 oThrSerGluLeuGlyArgLysGluAspThrLysSerThrSerAspSerGluIleIleSe 880

Qy       2821 TGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 rValSerAspThrGlnAsnTyrGluCysLeuProGluAlaThrTyrGlnLysGluIleLy 900

Qy       2881 GACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCTTCTCACTTTGAGCCTGCCAC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 sThrThrAsnGlyLysIleGluGluSerProGluLysProSerHisPheGluProAlaTh 920

Qy       2941 TGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAG 3000
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 rGluMetGlnAsnSerValProAsnLysGlyLeuGluTrpLysAsnLysGlnThrLeuAr 940

Qy       3001 AGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTGTGAAAGAGGAAG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 gAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeuProSerCysGluArgGlyAr 960

Qy       3061 GGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACAAATGAAAAATAA 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 gGluLeuLysLysAspAsnCysGluGlnIleThrAlaLysMetGluGlnMetLysAsnLy 980

Qy       3121 GTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATCACAGTTAGAGAA 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 sPheCysValLeuGlnLysGluLeuSerGluAlaLysGluIleLysSerGlnLeuGluAs 1000

Qy       3181 CCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTTAAATCAAGAAGA 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 nGlnLysAlaLysTrpGluGlnGluLeuCysSerValArgLeuThrLeuAsnGlnGluGl 1020

Qy       3241 AGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCCGAAGAGCAACTTAG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 uGluLysArgArgAsnValAspIleLeuLysGluLysIleArgProGluGluGlnLeuAr 1040

Qy       3301 GAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTCAGAATACAAGATATAGAATT 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 gLysLysLeuGluValLysGlnGlnLeuGluGlnThrLeuArgIleGlnAspIleGluLe 1060

Qy       3361 GAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACTCATGAAAGTGAAAATGATCT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 uLysSerValThrSerAsnLeuAsnGlnValSerHisThrHisGluSerGluAsnAspLe 1080

Qy       3421 CTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATGCTAAAACTGGAAGTAGCCAC 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 uPheHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluValAlaTh 1100

Qy       3481 ACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTTGAGGACATTAAGATTTTACA 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 rLeuLysHisGlnHisGlnValLysGluAsnLysTyrPheGluAspIleLysIleLeuGl 1120

Qy       3541 AGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAACAGAAAACAGTAACAAAAAG 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 nGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGlnLysThrValThrLysAr 1140

Qy       3601 GGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCAGAGAACACGATGCTGACTTC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 gAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAlaGluAsnThrMetLeuThrSe 1160

Qy       3661 TAAATTGAAGGAA 3673
             |||||||||||||
Db       1160 rLysLeuLysGlu 1164
```

RESULT 2
US-10-408-765A-1557
; Sequence 1557, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1557
; LENGTH: 1011
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-408-765A-1557

Alignment Scores:

Pred. No.:	3.16e-302	Length:	1011
Score:	5173.00	Matches:	1011
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.23%	Indels:	0
DB:	16	Gaps:	0

US-09-602-362E-26 (1-3673) x US-10-408-765A-1557 (1-1011)

```
Qy        641 ATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTC 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValAlaThrLeuLeuSerTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20

Qy        701 ACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACA 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrProLeuLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr 40

Qy        761 AAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATA 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysAsnAlaAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle 60

Qy        821 TGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCT 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAla 80

Qy        881 GAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATT 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GluAspIleHisGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100

Qy        941 CATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCA 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120

Qy       1001 GAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGAC 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluGlyThrSerThrGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAsp 140

Qy       1061 ACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACG 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160

Qy       1121 TCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAA 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180

Qy       1181 GAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerTrpPro 200

Qy       1241 GCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAA 1300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AlaLysGluArgSerArgLysIleThrTrpGluGluLysGluThrSerValLysThrGlu 220

Qy       1301 TGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATG 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240

Qy       1361 ATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCT 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSer 260

Qy       1421 GTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAA 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280

Qy       1481 GTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTAT 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300

Qy       1541 ACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320

Qy       1601 GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCT 1660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluGluTyrSerTrpAspSer 340

Qy       1661 GGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360

Qy       1721 AAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAG 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380

Qy       1781 CCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAA 1840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400

Qy       1841 ACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAAT 1900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAspGluGluAsn 420

Qy       1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA 1960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerTrpAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440

Qy       1961 GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAA 2020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluGluSerProValLys 460

Qy       2021 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480

Qy       2081 TTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCT 2140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500

Qy       2141 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520

Qy       2201 CTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540

Qy       2261 TCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaGlnMet 560

Qy       2321 TTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTC 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580

Qy       2381 CTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTC 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600

Qy       2441 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620

Qy       2501 TGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTC 2560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640

Qy       2561 AAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACT 2620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660

Qy       2621 GAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCA 2680
```

```
Db        661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680

Qy       2681 ACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAA 2740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGln 700

Qy       2741 GATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACT 2800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720

Qy       2801 TCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCT 2860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SerAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAla 740

Qy       2861 ACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCT 2920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 ThrTyrGlnLysGluIleLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760

Qy       2921 TCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGG 2980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780

Qy       2981 AAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTT 3040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800

Qy       3041 CCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAA 3100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820

Qy       3101 ATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAA 3160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840

Qy       3161 ATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA 3220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IleLysSerGlnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860

Qy       3221 TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT 3280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 LeuProLeuAsnGlnGluGluGluLysArgArgAsnValAspIleLeuLysGluLysIle 880

Qy       3281 AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTC 3340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGluGlnThrLeu 900

Qy       3341 AGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACT 3400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920

Qy       3401 CATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATG 3460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940

Qy       3461 CTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTT 3520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrPhe 960

Qy       3521 GAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAA 3580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLys 980

Qy       3581 CAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCA 3640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 GlnLysThrValThrLysArgAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAla 1000

Qy       3641 GAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              |||||||||||||||||||||||||||||||||
Db       1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011

RESULT 3
US-10-007-805-577

; Sequence 577, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577
;   LENGTH: 1239
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-007-805-577

Alignment Scores:
Pred. No.:                3.24e-302       Length:          1239
Score:                    5173.00         Matches:         1011
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.23%          Indels:          0
DB:                       13              Gaps:            0

US-09-602-362E-26 (1-3673) x US-10-007-805-577 (1-1239)

Qy        641 ATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTC 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValAlaThrLeuLeuSerTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20

Qy        701 ACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACA 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrProLeuLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr 40

Qy        761 AAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATA 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysAsnAlaAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle 60

Qy        821 TGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCT 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAla 80

Qy        881 GAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATT 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GluAspIleHisGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100

Qy        941 CATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCA 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120

Qy       1001 GAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGAC 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluGlyThrSerThrGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAsp 140

Qy       1061 ACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACG 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160

Qy       1121 TCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAA 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180
```

```
Qy    1181 GAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerTrpPro 200

Qy    1241 GCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAA 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     201 AlaLysGluArgSerArgLysIleThrTrpGluGluLysGluThrSerValLysThrGlu 220

Qy    1301 TGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATG 1360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240

Qy    1361 ATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCT 1420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSer 260

Qy    1421 GTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAA 1480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280

Qy    1481 GTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTAT 1540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300

Qy    1541 ACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA 1600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320

Qy    1601 GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCT 1660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluGluTyrSerTrpAspSer 340

Qy    1661 GGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360

Qy    1721 AAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAG 1780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 LysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380

Qy    1781 CCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAA 1840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400

Qy    1841 ACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAAT 1900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAspGluGluAsn 420

Qy    1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA 1960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 SerTrpAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440

Qy    1961 GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAA 2020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluSerProValLys 460

Qy    2021 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480

Qy    2081 TTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCT 2140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500

Qy    2141 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520

Qy    2201 CTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540

Qy    2261 TCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaAlaGlnMet 560

Qy    2321 TTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTC 2380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580

Qy    2381 CTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTC 2440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600

Qy    2441 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620

Qy    2501 TGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTC 2560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640

Qy    2561 AAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACT 2620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660

Qy    2621 GAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCA 2680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680

Qy    2681 ACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAA 2740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     681 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGln 700

Qy    2741 GATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACT 2800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720

Qy    2801 TCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCT 2860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 SerAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAla 740

Qy    2861 ACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCT 2920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     741 ThrTyrGlnLysGluIleLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760

Qy    2921 TCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGG 2980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780

Qy    2981 AAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTT 3040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800

Qy    3041 CCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAA 3100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820

Qy    3101 ATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAA 3160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840

Qy    3161 ATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA 3220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 IleLysSerGlnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860

Qy    3221 TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT 3280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     861 LeuProLeuAsnGlnGluGluGluLysArgArgAsnValAspIleLeuLysGluLysIle 880

Qy    3281 AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTC 3340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGluGlnThrLeu 900

Qy    3341 AGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACT 3400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920

Qy       3401 CATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATG 3460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940

Qy       3461 CTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTT 3520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrPhe 960

Qy       3521 GAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAA 3580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLys 980

Qy       3581 CAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCA 3640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 GlnLysThrValThrLysArgAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAla 1000

Qy       3641 GAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              |||||||||||||||||||||||||||||||||
Db       1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011

RESULT 4
US-10-076-622-577
; Sequence 577, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C11
;  CURRENT APPLICATION NUMBER: US/10/076,622
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577
;   LENGTH: 1239
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-076-622-577

Alignment Scores:
Pred. No.:                3.24e-302       Length:          1239
Score:                    5173.00         Matches:         1011
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.23%          Indels:          0
DB:                       14              Gaps:            0

US-09-602-362E-26 (1-3673) x US-10-076-622-577 (1-1239)

Qy        641 ATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTC 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValAlaThrLeuLeuSerTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20

Qy        701 ACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACA 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrProLeuLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr 40

Qy        761 AAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATA 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysAsnAlaAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle 60

Qy        821 TGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCT 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAla 80

Qy        881 GAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATT 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GluAspIleHisGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100

Qy        941 CATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCA 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120

Qy       1001 GAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGAC 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluGlyThrSerThrGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAsp 140

Qy       1061 ACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACG 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160

Qy       1121 TCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAA 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180

Qy       1181 GAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerTrpPro 200

Qy       1241 GCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAA 1300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AlaLysGluArgSerArgLysIleThrTrpGluGluLysGluThrSerValLysThrGlu 220

Qy       1301 TGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATG 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240

Qy       1361 ATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCT 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSer 260

Qy       1421 GTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAA 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280

Qy       1481 GTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTAT 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300

Qy       1541 ACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320

Qy       1601 GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCT 1660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluGluTyrSerTrpAspSer 340

Qy       1661 GGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360

Qy       1721 AAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAG 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380

Qy       1781 CCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAA 1840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400

Qy       1841 ACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAAT 1900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAspGluGluAsn 420

Qy       1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA 1960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerTrpAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440

Qy       1961 GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAA 2020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluGluSerProValLys 460

Qy       2021 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2080
```

```
Db        461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480

Qy       2081 TTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCT 2140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500

Qy       2141 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520

Qy       2201 CTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540

Qy       2261 TCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaAlaGlnMet 560

Qy       2321 TTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTC 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580

Qy       2381 CTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTC 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600

Qy       2441 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620

Qy       2501 TGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTC 2560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640

Qy       2561 AAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACT 2620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660

Qy       2621 GAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCA 2680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680

Qy       2681 ACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAA 2740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGln 700

Qy       2741 GATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACT 2800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720

Qy       2801 TCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCT 2860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SerAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAla 740

Qy       2861 ACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCT 2920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 ThrTyrGlnLysGluIleLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760

Qy       2921 TCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGG 2980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780

Qy       2981 AAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTT 3040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800

Qy       3041 CCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAA 3100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820

Qy       3101 ATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAA 3160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840

Qy       3161 ATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA 3220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IleLysSerGlnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860

Qy       3221 TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT 3280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 LeuProLeuAsnGlnGluGluGluLysArgArgAsnValAspIleLeuLysGluLysIle 880

Qy       3281 AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTC 3340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGluGlnThrLeu 900

Qy       3341 AGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACT 3400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920

Qy       3401 CATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATG 3460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940

Qy       3461 CTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTT 3520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrPhe 960

Qy       3521 GAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAA 3580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLys 980

Qy       3581 CAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCA 3640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 GlnLysThrValThrLysArgAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAla 1000

Qy       3641 GAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              |||||||||||||||||||||||||||||||||
Db       1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011
```

RESULT 5
US-10-124-805-577
; Sequence 577, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-577

Alignment Scores:

Pred. No.:	3.24e-302	Length:	1239
Score:	5173.00	Matches:	1011
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.23%	Indels:	0
DB:	14	Gaps:	0

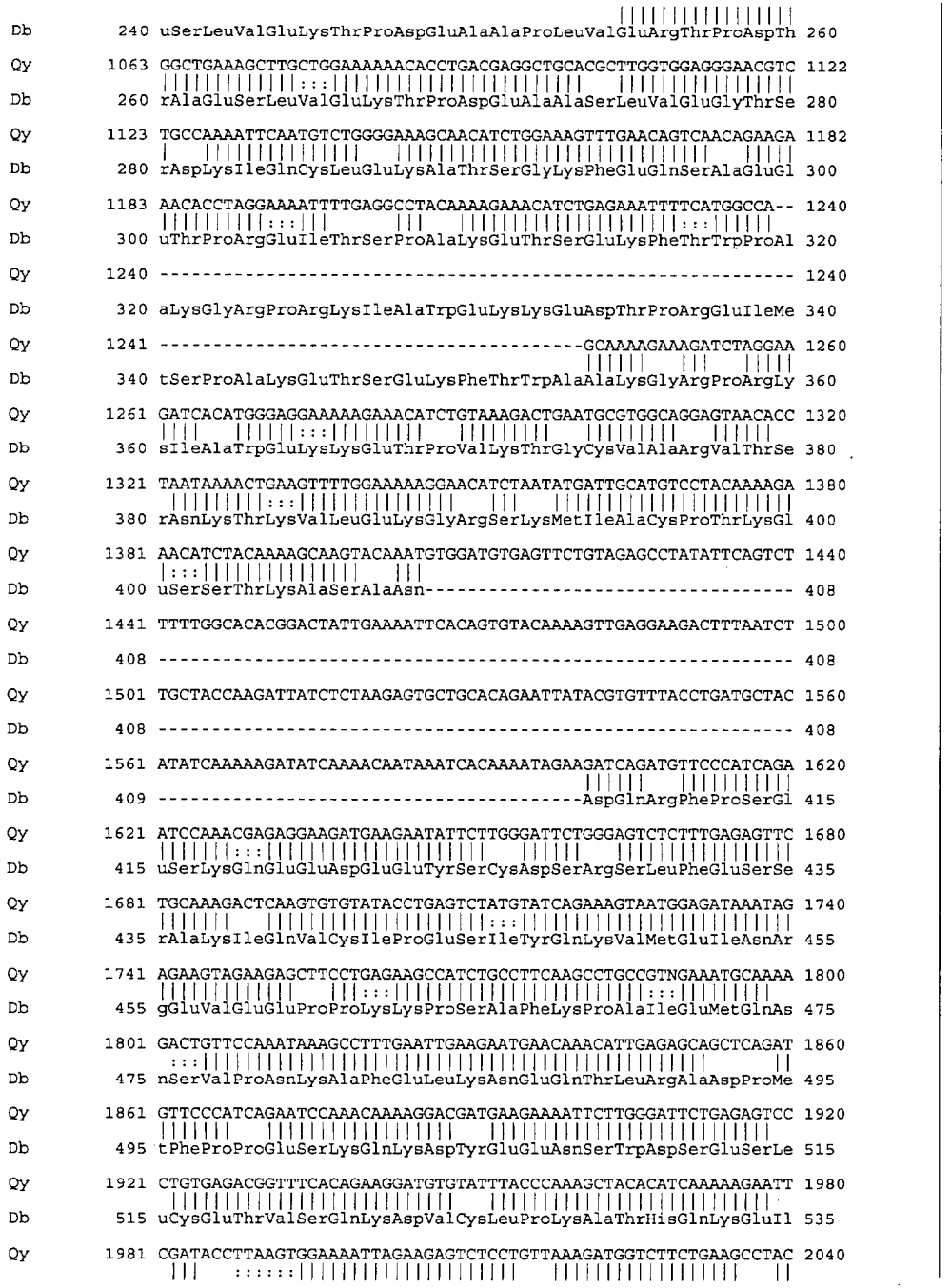
US-09-602-362E-26 (1-3673) x US-10-124-805-577 (1-1239)

```
Qy        641 ATGGTGGCAACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTC 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValAlaThrLeuLeuSerTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20
```

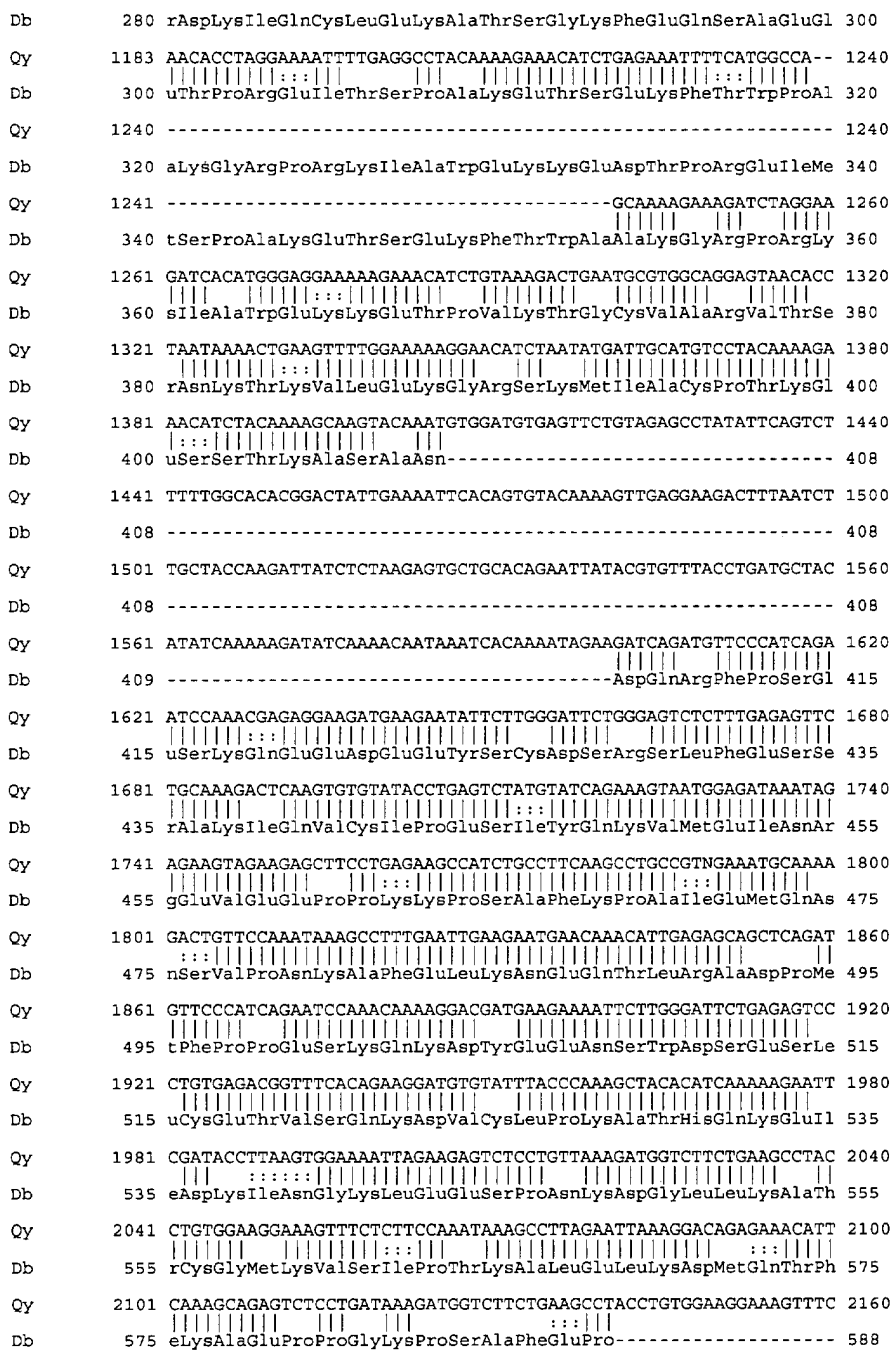
```
Qy       701 ACACCCCTTTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACA 760
Db        21 ThrProLeuLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr  40

Qy       761 AAAAATGCAAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATA 820
Db        41 LysAsnAlaAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle  60

Qy       821 TGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCT 880
Db        61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAla  80

Qy       881 GAAGACATACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATT 940
Db        81 GluAspIleHisGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100

Qy       941 CATCAACAACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCA 1000
Db       101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120

Qy      1001 GAAGGAACATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCGGAAAGAACACCTGAC 1060
Db       121 GluGlyThrSerThrGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAsp 140

Qy      1061 ACGGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACG 1120
Db       141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160

Qy      1121 TCTGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAA 1180
Db       161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180

Qy      1181 GAAACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA 1240
Db       181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerTrpPro 200

Qy      1241 GCAAAAGAAAGATCTAGGAAGATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAA 1300
Db       201 AlaLysGluArgSerArgLysIleThrTrpGluGluLysGluThrSerValLysThrGlu 220

Qy      1301 TGCGTGGCAGGAGTAACACCTAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATG 1360
Db       221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240

Qy      1361 ATTGCATGTCCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCT 1420
Db       241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSer 260

Qy      1421 GTAGAGCCTATATTCAGTCTTTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAA 1480
Db       261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280

Qy      1481 GTTGAGGAAGACTTTAATCTTGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTAT 1540
Db       281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300

Qy      1541 ACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAA 1600
Db       301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320

Qy      1601 GATCAGATGTTCCCATCAGAATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCT 1660
Db       321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluGluTyrSerTrpAspSer 340

Qy      1661 GGGAGTCTCTTTGAGAGTTCTGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720
Db       341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360

Qy      1721 AAAGTAATGGAGATAAATAGAGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAG 1780
Db       361 LysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380

Qy      1781 CCTGCCGTNGAAATGCAAAAGACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAA 1840
Db       381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400

Qy      1841 ACATTGAGAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAAT 1900
Db       401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspAspGluGluAsn 420

Qy      1901 TCTTGGGATTCTGAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAA 1960
Db       421 SerTrpAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440

Qy      1961 GCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAA 2020
Db       441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluGluSerProValLys 460

Qy      2021 GATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAA 2080
Db       461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480

Qy      2081 TTAAAGGACAGAGAAACATTCAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCT 2140
Db       481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500

Qy      2141 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACA 2200
Db       501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520

Qy      2201 CTCAAAGCAGAGTCTCCTGATAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260
Db       521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540

Qy      2261 TCTCTTCCAAATAAAGCTTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320
Db       541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaAlaGlnMet 560

Qy      2321 TTCCCATCAGAATCCAAACAAAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTC 2380
Db       561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580

Qy      2381 CTTGAGACTCTCTTACAGAATGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTC 2440
Db       581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600

Qy      2441 GATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACC 2500
Db       601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620

Qy      2501 TGTGGAATGAAAATTTCTCTTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTC 2560
Db       621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640

Qy      2561 AAAGCAGAGGATGTGAGTTCTGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACT 2620
Db       641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660

Qy      2621 GAAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCA 2680
Db       661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680

Qy      2681 ACAAAGACAGTAACTGGACAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAA 2740
Db       681 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGln 700

Qy      2741 GATCAAACAAATAAGATGCCCACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACT 2800
Db       701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720

Qy      2801 TCAGATTCTGAGATTATCTCTGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCT 2860
Db       721 SerAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAla 740

Qy      2861 ACATATCAAAAAGAAATAAAGACAACAAATGGCAAAATAGAAGAGTCTCCTGAAAAGCCT 2920
```

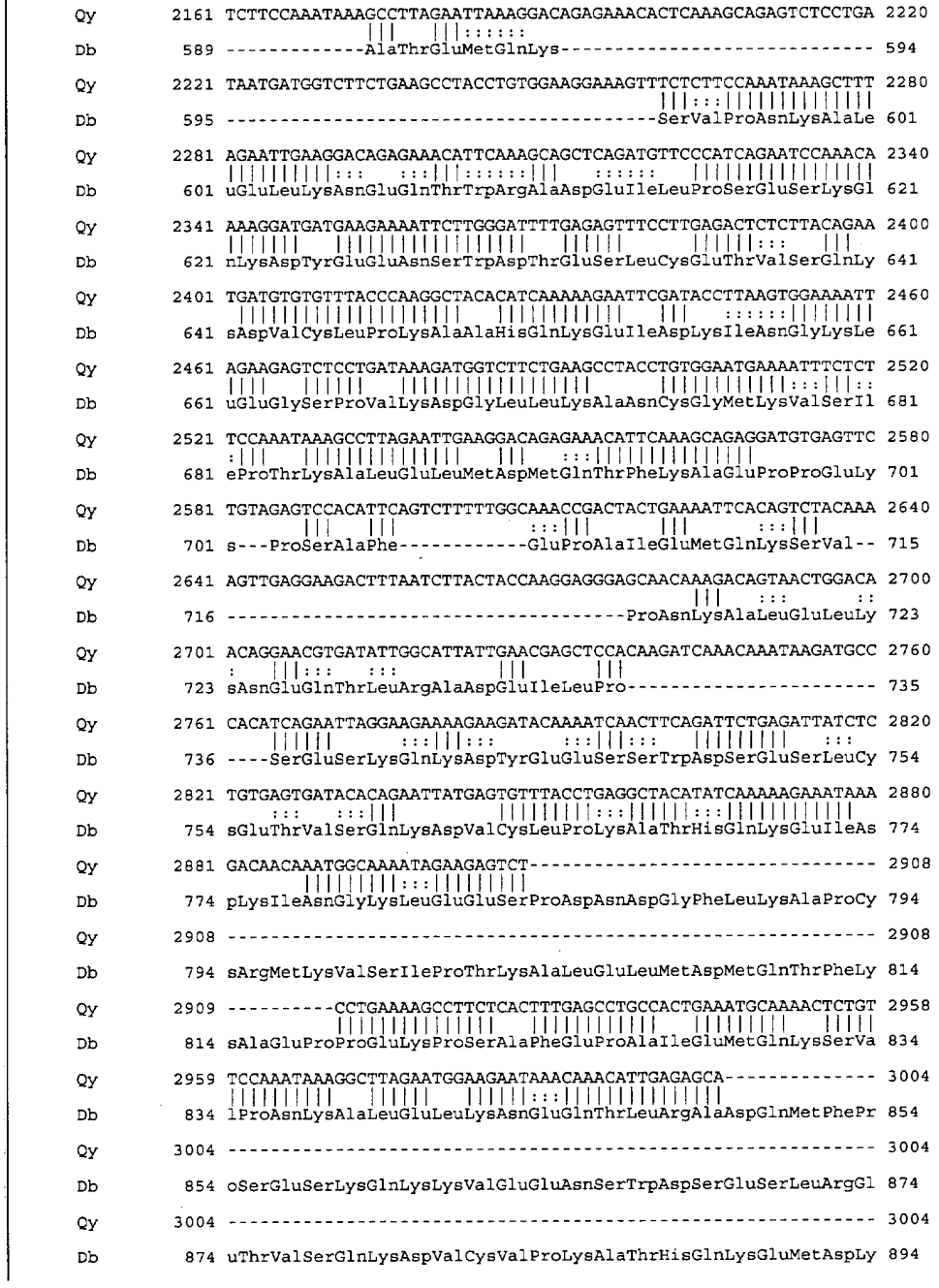
```
Db        741 ThrTyrGlnLysGluIleLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760

Qy       2921 TCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCCAAATAAAGGCTTAGAATGG 2980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780

Qy       2981 AAGAATAAACAAACATTGAGAGCAGATTCAACTACCCTATCAAAAATCTTGGATGCACTT 3040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800

Qy       3041 CCTTCTTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAA 3100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820

Qy       3101 ATGGAACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAA 3160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840

Qy       3161 ATAAAATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGA 3220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 IleLysSerGlnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860

Qy       3221 TTGCCTTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATT 3280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 LeuProLeuAsnGlnGluGluGluLysArgArgAsnValAspIleLeuLysGluLysIle 880

Qy       3281 AGACCCGAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACAGACTCTC 3340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGluGlnThrLeu 900

Qy       3341 AGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTCTCACACT 3400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920

Qy       3401 CATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAATTGCCATG 3460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940

Qy       3461 CTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAAATACTTT 3520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrPhe 960

Qy       3521 GAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAAACTGAAA 3580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLys 980

Qy       3581 CAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCTGACGGCA 3640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 GlnLysThrValThrLysArgAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAla 1000

Qy       3641 GAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              |||||||||||||||||||||||||||||||||
Db       1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011

RESULT 6
US-10-058-270A-4
; Sequence 4, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Afar, Daniel
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Breast Cancer
;  FILE REFERENCE: 018501-005210US
;  CURRENT APPLICATION NUMBER: US/10/058,270A
;  CURRENT FILING DATE:  2002-01-24
;  PRIOR APPLICATION NUMBER: US 60/263,965
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: US 60/265,928
;  PRIOR FILING DATE: 2001-02-02
;  PRIOR APPLICATION NUMBER: US 09/829,472
;  PRIOR FILING DATE: 2001-04-09
;  PRIOR APPLICATION NUMBER: US 60/282,698
;  PRIOR FILING DATE: 2001-04-09
;  PRIOR APPLICATION NUMBER: US 60/288,590
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/294,443
;  PRIOR FILING DATE: 2001-05-29
;  NUMBER OF SEQ ID NOS: 141
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 1341
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-058-270A-4

Alignment Scores:
Pred. No.:                2.71e-199      Length:          1341
Score:                    3458.00        Matches:         740
Percent Similarity:       65.58%         Conservative:    83
Best Local Similarity:    58.96%         Mismatches:      159
Query Match:              53.63%         Indels:          273
DB:                       12             Gaps:            12

US-09-602-362E-26 (1-3673) x US-10-058-270A-4 (1-1341)

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
              ||||||     ||||||   :::|||||||||   :::|||   :::||||||||||||
Db          1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
              ||||||||||||||||||||||||||      ||||||||||||||||||||||||||||
Db         21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
              |||   |||:::|||||||||||||||   |||||||||||||||||||||||||||||
Db         41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
                 :::||||||  ||||||||||||||||||||:::|||||||||:::|||   |||||
Db         61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
              ||||||||||||   ||||||||||||||||||||||||||   |||   :::|||||
Db         80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
              |   |||||||||:::|||||||||||||||||   ||||||||||||||||||||||
Db        100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
              |||||||:::|||   ||||||||||:::|||   |||||||||||||   ||||||||
Db        120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
              ||||||||||||   |||:::   |||||||||||||||||||||:::|||   ||
Db        140 aAsnAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGl 160

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIl 180

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
              |   |||:::|||||||||   ||||||         |||   ::::::||||||:::||
Db        180 eCysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGl 200

Qy        949 ACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAAC 1008
              |::::::||:::|||||||||||   ||||||   ||||||||||||||||||||||
Db        200 nIleMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyTh 220

Qy       1009 ATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCG---------------------- 1045
              ||||   ||||||||||||||||||||||||||||
Db        220 rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGl 240
```

```
Db        834 lProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854

Qy       3004 ------------------------------------------------------------ 3004

Db        854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874

Qy       3004 ------------------------------------------------------------ 3004

Db        874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894

Qy       3005 -------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTC 3045

Db        894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914

Qy       3046 TTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGA 3105

Db        914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934

Qy       3106 ACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAA 3165

Db        934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954

Qy       3166 ATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCC 3225

Db        954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974

Qy       3226 TTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285

Db        974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGl 994

Qy       3286 C---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGA 3330

Db        994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014

Qy       3331 ACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGT 3390

Db       1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034

Qy       3391 TTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGA 3450

Db       1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

Qy       3451 AATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAA 3510

Db       1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074

Qy       3511 TAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCT 3570

Db       1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094

Qy       3571 AAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGT 3630

Db       1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114

Qy       3631 TCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673

Db       1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128

RESULT 7
US-10-007-805-565
; Sequence 565, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 565
;   LENGTH: 1341
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-007-805-565

Alignment Scores:
Pred. No.:              2.71e-199       Length:         1341
Score:                  3458.00         Matches:        740
Percent Similarity:     65.58%          Conservative:   83
Best Local Similarity:  58.96%          Mismatches:     159
Query Match:            53.63%          Indels:         273
DB:                     13              Gaps:           12

US-09-602-362E-26 (1-3673) x US-10-007-805-565 (1-1341)

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408

Db          1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468

Db         21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528

Db         41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588

Db         61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648

Db         80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708

Db        100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768

Db        120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828

Db        140 aAsnAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGl 160

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888

Db        160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIl 180

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948

Db        180 eCysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGl 200

Qy        949 ACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAAC 1008

Db        200 nIleMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyTh 220

Qy       1009 ATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCG---------------------- 1045

Db        220 rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGl 240

Qy       1046 ----------------------------------------GAAAGAACACCTGACAC 1062
```


Db 240 uSerLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspTh 260
Qy 1063 GGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTC 1122
Db 260 rAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSe 280
Qy 1123 TGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGA 1182
Db 280 rAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGl 300
Qy 1183 AACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA-- 1240
Db 300 uThrProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAl 320
Qy 1240 ---------- 1240
Db 320 aLysGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMe 340
Qy 1241 ----------GCAAAAGAAAGATCTAGGAA 1260
Db 340 tSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLy 360
Qy 1261 GATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACC 1320
Db 360 sIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSe 380
Qy 1321 TAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGA 1380
Db 380 rAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGl 400
Qy 1381 AACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCT 1440
Db 400 uSerSerThrLysAlaSerAlaAsn---------- 408
Qy 1441 TTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCT 1500
Db 408 ---------- 408
Qy 1501 TGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTAC 1560
Db 408 ---------- 408
Qy 1561 ATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGA 1620
Db 409 ----------AspGlnArgPheProSerGl 415
Qy 1621 ATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTC 1680
Db 415 uSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSe 435
Qy 1681 TGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAG 1740
Db 435 rAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnAr 455
Qy 1741 AGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAA 1800
Db 455 gGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAs 475
Qy 1801 GACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGAT 1860
Db 475 nSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMe 495
Qy 1861 GTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCC 1920
Db 495 tPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLe 515
Qy 1921 CTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATT 1980
Db 515 uCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIl 535
Qy 1981 CGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTAC 2040
Db 535 eAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaTh 555
Qy 2041 CTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATT 2100
Db 555 rCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPh 575
Qy 2101 CAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTC 2160
Db 575 eLysAlaGluProProGlyLysProSerAlaPheGluPro---------- 588
Qy 2161 TCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGA 2220
Db 589 ------------AlaThrGluMetGlnLys---------- 594
Qy 2221 TAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTT 2280
Db 595 ----------SerValProAsnLysAlaLe 601
Qy 2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACA 2340
Db 601 uGluLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysGl 621
Qy 2341 AAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAA 2400
Db 621 nLysAspTyrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLy 641
Qy 2401 TGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATT 2460
Db 641 sAspValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLe 661
Qy 2461 AGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCT 2520
Db 661 uGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIl 681
Qy 2521 TCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTC 2580
Db 681 eProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLy 701
Qy 2581 TGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAA 2640
Db 701 s---ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal-- 715
Qy 2641 AGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACA 2700
Db 716 ----------ProAsnLysAlaLeuGluLeuLy 723
Qy 2701 ACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCC 2760
Db 723 sAsnGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 735
Qy 2761 CACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTC 2820
Db 736 ----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCy 754
Qy 2821 TGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAA 2880
Db 754 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 774
Qy 2881 GACAACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 774 pLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCy 794
Qy 2908 ---------- 2908
Db 794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLy 814
Qy 2909 ----------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958
Db 814 sAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834
Qy 2959 TCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 834 lProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854

```
Qy       3004 ------------------------------------------------------------ 3004

Db        854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874

Qy       3004 ------------------------------------------------------------ 3004

Db        874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894

Qy       3005 ------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTC 3045
                                |||||||||:::|||||||||||||||||||   :::   ||
Db        894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914

Qy       3046 TTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGA 3105
              ||||||||||   |||||||||:::||||||:::|||||||||   |||   ||||||||
Db        914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934

Qy       3106 ACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAA 3165
              ||||||||||   ||||||||||||||:::|||:::||||||||||||||||||||||||
Db        934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954

Qy       3166 ATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCC 3225
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974

Qy       3226 TTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285
              ||||||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGl 994

Qy       3286 C---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGA 3330
                             |||||||||   ||||||:::|||||||||||   ||||||||
Db        994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014

Qy       3331 ACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGT 3390
              ||||   ||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034

Qy       3391 TTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGA 3450
              ||||||||||||||:::||||||   |||   ||||||||||||||||||||||||||||
Db       1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

Qy       3451 AATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAA 3510
              ||||||||||||||||||||||:::|||||||||||||||||:::|||   ||||||||
Db       1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074

Qy       3511 TAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCT 3570
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db       1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094

Qy       3571 AAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGT 3630
              ||||||||||:::::::::::::|||||||||||||||||||||      ||||||||||
Db       1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114

Qy       3631 TCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              ||||   ||||||||||||||||||||||||||||||||||
Db       1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128

RESULT 8
US-10-076-622-565
; Sequence 565, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C11
;  CURRENT APPLICATION NUMBER: US/10/076,622
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 565
;  LENGTH: 1341
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-076-622-565

Alignment Scores:
Pred. No.:                2.71e-199      Length:          1341
Score:                    3458.00        Matches:         740
Percent Similarity:       65.58%         Conservative:    83
Best Local Similarity:    58.96%         Mismatches:      159
Query Match:              53.63%         Indels:          273
DB:                       14             Gaps:            12

US-09-602-362E-26 (1-3673) x US-10-076-622-565 (1-1341)

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
              ||||||      ||||||   :::|||||||||   :::|||   :::||||||||||||
Db          1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
              ||||||||||||||||||||||||||      ||||||||||||||||||||||||||||
Db         21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
              |||   |||:::|||||||||||||   |||||||||||||||||||||||||||||||
Db         41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
                 :::|||||| ||||||||||||||||||:::|||||||||:::|||   |||||
Db         61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
              ||||||||||||   ||||||||||||||||||||||||||||   |||   :::|||||
Db         80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
              |   |||||||||:::|||||||||||||||||||   |||||||||||||||||||||
Db        100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
              |||||||:::|||   |||||||||:::|||   ||||||||||||   ||||||||
Db        120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuLeuLysAsnAl 140

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
              ||||||||||||   |||:::   ||||||||||||||||||||||||:::|||   ||
Db        140 aAsnAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGl 160

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIl 180

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
              |   |||:::|||||||||   ||||||          |||   ::::::|||||:::||
Db        180 eCysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGl 200

Qy        949 ACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAAC 1008
              |::::::|||:::|||||||||||   ||||||   ||||||||||||||||||||||
Db        200 nIleMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyTh 220

Qy       1009 ATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCG---------------------- 1045
              ||||   ||||||||||||||||||||||||||||||
Db        220 rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGl 240

Qy       1046 ----------------------------------------GAAAGAACACCTGACAC 1062
                                                     |||||||||||||||||
Db        240 uSerLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspTh 260

Qy       1063 GGCTGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTC 1122
              ||||||||||||:::|||||||||||||||||||||||||   |||||||||||||||||
Db        260 rAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSe 280

Qy       1123 TGCCAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGA 1182
              |   ||||||||||||||   |||||||||||||||||||||||||||||   |||||
```


Db 280 rAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGl 300
Qy 1183 AACACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA-- 1240
Db 300 uThrProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAl 320
Qy 1240 ---------- 1240
Db 320 aLysGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMe 340
Qy 1241 ----------GCAAAAGAAAGATCTAGGAA 1260
Db 340 tSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLy 360
Qy 1261 GATCACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACC 1320
Db 360 sIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSe 380
Qy 1321 TAATAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGA 1380
Db 380 rAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGl 400
Qy 1381 AACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCT 1440
Db 400 uSerSerThrLysAlaSerAlaAsn---------- 408
Qy 1441 TTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCT 1500
Db 408 ---------- 408
Qy 1501 TGCTACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTAC 1560
Db 408 ---------- 408
Qy 1561 ATATCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGA 1620
Db 409 ----------AspGlnArgPheProSerGl 415
Qy 1621 ATCCAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTC 1680
Db 415 uSerLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSe 435
Qy 1681 TGCAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAG 1740
Db 435 rAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnAr 455
Qy 1741 AGAAGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAA 1800
Db 455 gGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAs 475
Qy 1801 GACTGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGAT 1860
Db 475 nSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMe 495
Qy 1861 GTTCCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCC 1920
Db 495 tPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLe 515
Qy 1921 CTGTGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATT 1980
Db 515 uCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIl 535
Qy 1981 CGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTAC 2040
Db 535 eAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaTh 555
Qy 2041 CTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATT 2100
Db 555 rCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPh 575
Qy 2101 CAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTC 2160
Db 575 eLysAlaGluProProGlyLysProSerAlaPheGluPro---------- 588

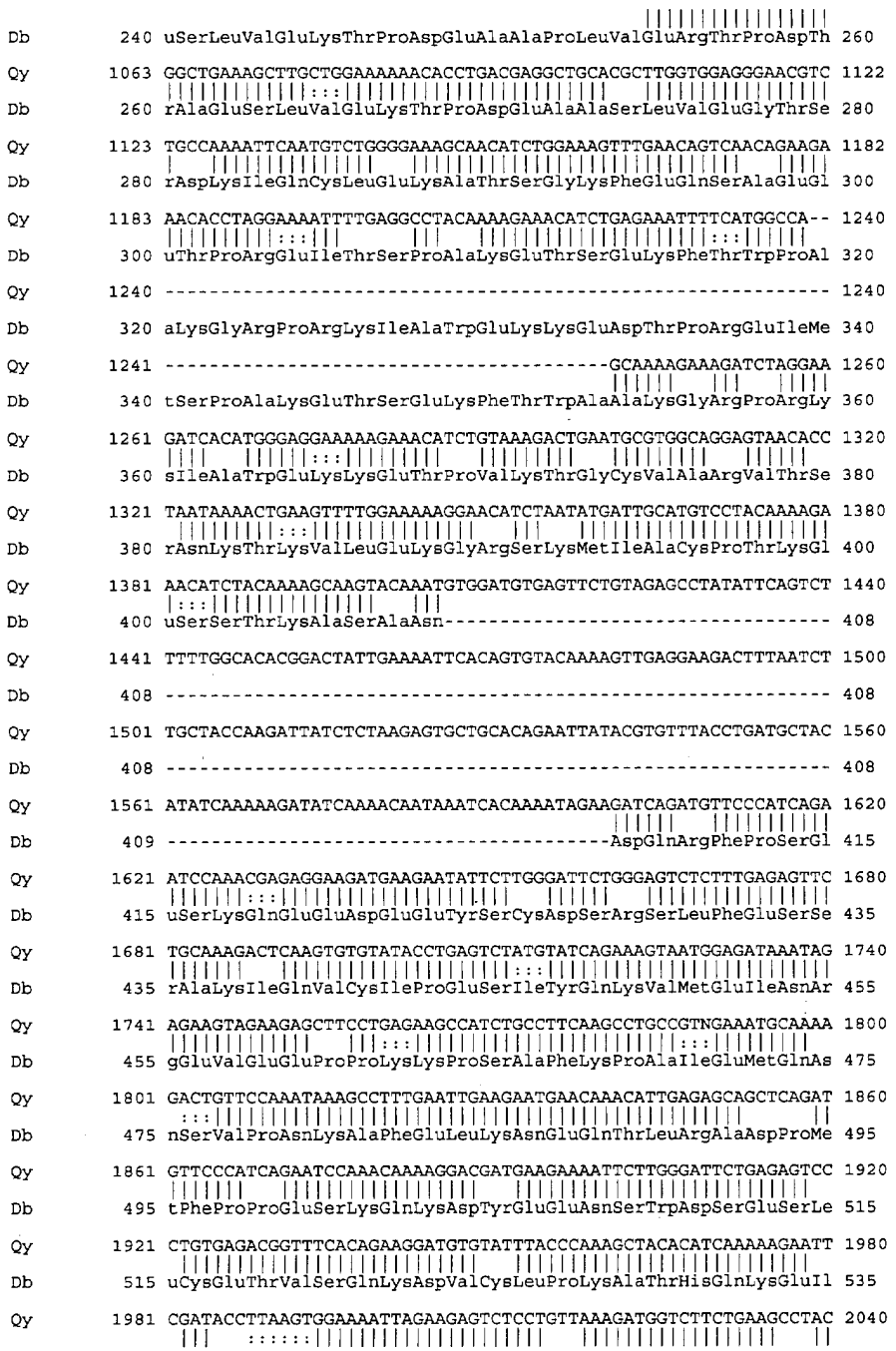

Qy 2161 TCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGA 2220
Db 589 ----------AlaThrGluMetGlnLys---------- 594
Qy 2221 TAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTT 2280
Db 595 ----------SerValProAsnLysAlaLe 601
Qy 2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACA 2340
Db 601 uGluLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysGl 621
Qy 2341 AAAGGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAA 2400
Db 621 nLysAspTyrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLy 641
Qy 2401 TGATGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATT 2460
Db 641 sAspValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLe 661
Qy 2461 AGAAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCT 2520
Db 661 uGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIl 681
Qy 2521 TCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTC 2580
Db 681 eProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLy 701
Qy 2581 TGTAGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAA 2640
Db 701 s---ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal-- 715
Qy 2641 AGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACA 2700
Db 716 ----------ProAsnLysAlaLeuGluLeuLy 723
Qy 2701 ACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCC 2760
Db 723 sAsnGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 735
Qy 2761 CACATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTC 2820
Db 736 ----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCy 754
Qy 2821 TGTGAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAA 2880
Db 754 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 774
Qy 2881 GACAACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 774 pLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCy 794
Qy 2908 ---------- 2908
Db 794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLy 814
Qy 2909 ----------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958
Db 814 sAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834
Qy 2959 TCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 834 lProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854
Qy 3004 ---------- 3004
Db 854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874
Qy 3004 ---------- 3004
Db 874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894

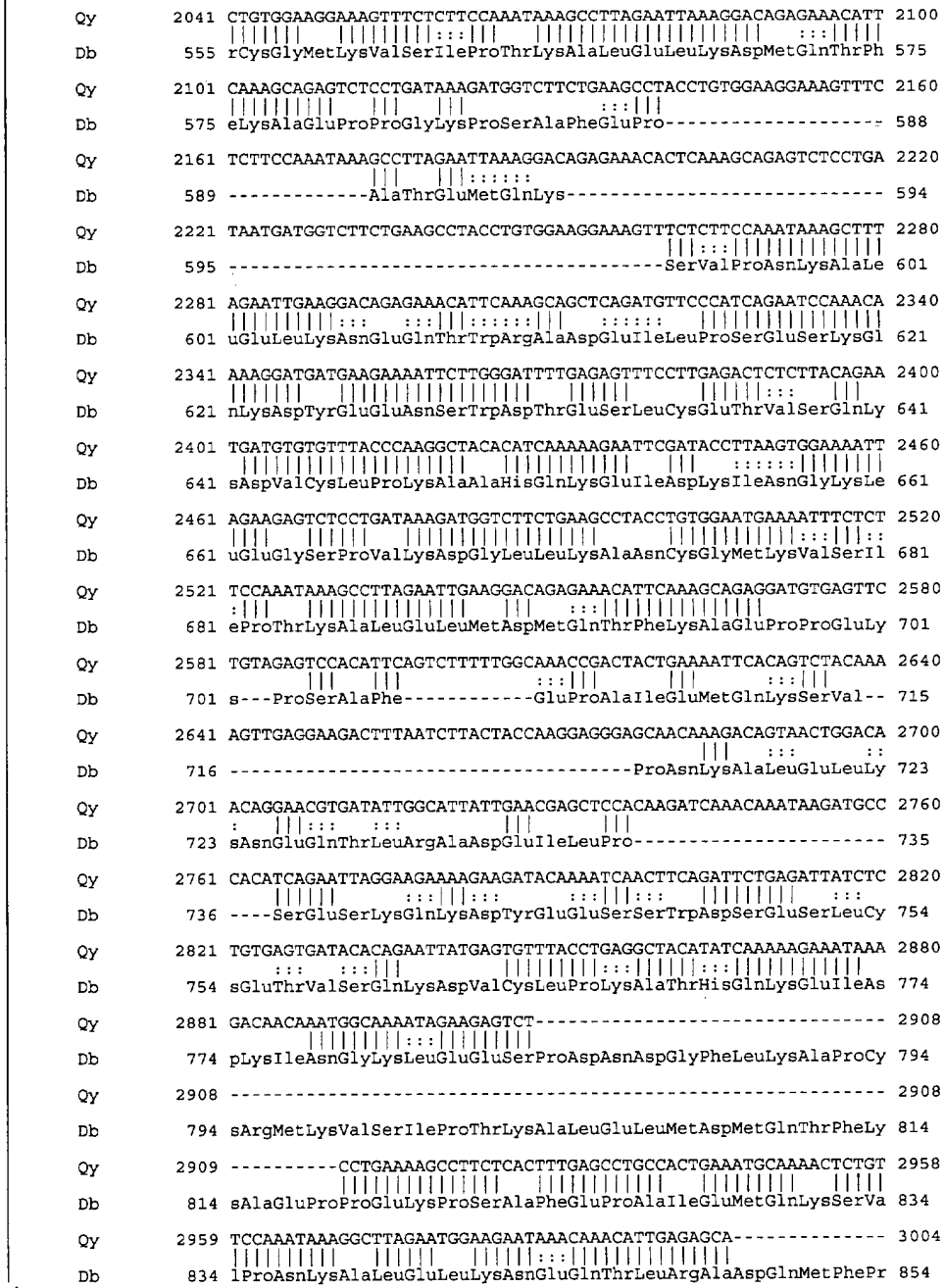
```
Qy       3005 ------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTC 3045
                                ||||||||:::||||||||||||||||     :::   .||
Db        894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914

Qy       3046 TTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGA 3105
              |||||||||   ||||||||:::|||||:::||||||||   |||   |||||||
Db        914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934

Qy       3106 ACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAA 3165
              |||||||||   ||||||||||||||:::|||:::||||||||||||||||||||
Db        934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954

Qy       3166 ATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCC 3225
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974

Qy       3226 TTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285
              ||||||||||||||||||||||||||||||||   |||||||||   ||||||||||
Db        974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGl 994

Qy       3286 C--------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGA 3330
                             ||||||||   |||||:::||||||||||||   |||||||
Db        994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014

Qy       3331 ACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGT 3390
              ||||  ||||||||||||||||||||||||||||||||   |||||||||||||||
Db       1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034

Qy       3391 TTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGA 3450
              ||||||||||||||||:::||||||   |||   |||||||||||||||||||||||||
Db       1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

Qy       3451 AATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAA 3510
              ||||||||||||||||||||||:::|||||||||||||||||||:::|||   |||||||
Db       1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074

Qy       3511 TAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCT 3570
              |||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db       1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094

Qy       3571 AAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGT 3630
              |||||||||::::::::::::|||||||||||||||||||||     ||||||||||
Db       1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114

Qy       3631 TCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              ||||  ||||||||||||||||||||||||||||||||||||
Db       1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128
```

RESULT 9
US-10-177-293-334
; Sequence 334, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 334
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-334

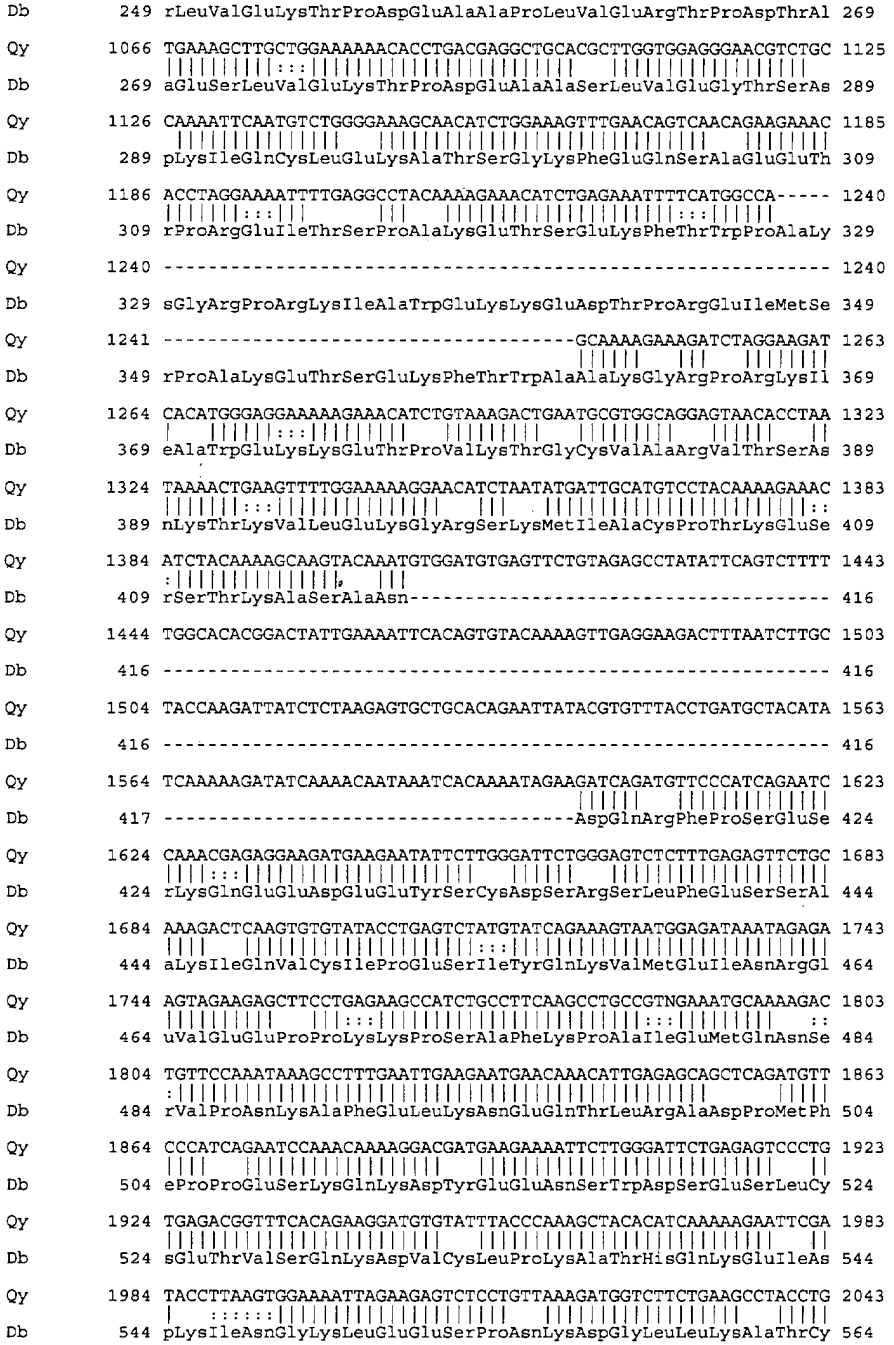
```
Alignment Scores:
Pred. No.:               2.71e-199      Length:          1341
Score:                   3458.00        Matches:         740
Percent Similarity:      65.58%         Conservative:    83
Best Local Similarity:   58.96%         Mismatches:      159
Query Match:             53.63%         Indels:          273
DB:                      14             Gaps:            12

US-09-602-362E-26 (1-3673) x US-10-177-293-334 (1-1341)

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
              |||||     |||||    :::||||||||    :::|||    :::|||||||||||
Db          1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
              ||||||||||||||||||||||||||||           ||||||||||||||||||||||
Db         21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
              |||   |||:::||||||||||||||   |||||||||||||||||||||||||||||
Db         41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
                 :::|||||| |||||||||||||||||||:::||||||||:::|||   |||||
Db         61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
              |||||||||||||   |||||||||||||||||||||||||||||||   |||   :::|||||
Db         80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
              |   ||||||||:::||||||||||||||||||   |||||||||||||||||||||
Db        100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
              ||||||:::|||   |||||||||:::|||   ||||||||||||||   ||||||||
Db        120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
              |||||||||||||   |||:::   |||||||||||||||||||||||||:::|||   ||
Db        140 aAsnAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGl 160

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIl 180

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
              |   |||:::|||||||||   ||||||          |||    ::::::||||||:::||
Db        180 eCysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGl 200
```

```
Db        794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLy 814

Qy       2909 ----------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958
                      ||||||||||||||||   ||||||||||||   |||||||||   |||||
Db        814 sAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834

Qy       2959 TCCAAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA-------------- 3004
              ||||||||||    ||||||   ||||||:::|||||||||||||||
Db        834 lProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854

Qy       3004 ------------------------------------------------------------ 3004

Db        854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874

Qy       3004 ------------------------------------------------------------ 3004

Db        874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894

Qy       3005 ------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTC 3045
                               |||||||||:::|||||||||||||||||   :::   ||
Db        894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914

Qy       3046 TTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGA 3105
              ||||||||||   |||||||||:::||||||:::|||||||||    |||   ||||||||
Db        914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934

Qy       3106 ACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAA 3165
              ||||||||||   |||||||||||||||:::|||:::||||||||||||||||||||||
Db        934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954

Qy       3166 ATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCC 3225
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974

Qy       3226 TTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285
                ||||||||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGl 994

Qy       3286 C---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGA 3330
                            |||||||||   ||||||:::|||||||||||   ||||||||
Db        994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014

Qy       3331 ACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGT 3390
              ||||   ||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034

Qy       3391 TTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGA 3450
              ||||||||||||||||:::||||||   |||   ||||||||||||||||||||||||||
Db       1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

Qy       3451 AATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAA 3510
              ||||||||||||||||||||||:::||||||||||||||||:::|||   ||||||||
Db       1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074

Qy       3511 TAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCT 3570
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db       1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094

Qy       3571 AAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGT 3630
              ||||||||||:::::::::::|||||||||||||||||||||   ||||||||||
Db       1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114

Qy       3631 TCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              ||||   ||||||||||||||||||||||||||||||||||||
Db       1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128

RESULT 10
US-10-124-805-565
; Sequence 565, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C12
;  CURRENT APPLICATION NUMBER: US/10/124,805
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 565
;   LENGTH: 1341
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-124-805-565

Alignment Scores:
Pred. No.:                2.71e-199      Length:          1341
Score:                    3458.00        Matches:         740
Percent Similarity:       65.58%         Conservative:    83
Best Local Similarity:    58.96%         Mismatches:      159
Query Match:              53.63%         Indels:          273
DB:                       14             Gaps:            12

US-09-602-362E-26 (1-3673) x US-10-124-805-565 (1-1341)

Qy        349 ATGACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCT 408
              ||||||      ||||||   :::|||||||||   :::|||   :::||||||||||||
Db          1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20

Qy        409 CTACACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAG 468
              ||||||||||||||||||||||||||||      ||||||||||||||||||||||||||
Db         21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40

Qy        469 TGCCNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGC 528
              |||   |||:::|||||||||||||||   ||||||||||||||||||||||||||||||
Db         41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60

Qy        529 GAGAGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGA 588
              :::|||||| ||||||||||||||||||||:::|||||||:::|||   |||||
Db         61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

Qy        589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGC 648
              ||||||||||||   ||||||||||||||||||||||||||   |||   :::|||||
Db         80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100

Qy        649 AACACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCT 708
              |   ||||||||:::|||||||||||||||||||||||||||||||||||||||||||
Db        100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120

Qy        709 TTTACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGC 768
              |||||||:::|||   |||||||||:::|||   ||||||||||||||   |||||||
Db        120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140

Qy        769 AAATGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGG 828
              ||||||||||||   |||:::   ||||||||||||||||||||:::|||   ||
Db        140 aAsnAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGl 160

Qy        829 CTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACAT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIl 180

Qy        889 ACATGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACA 948
              |   |||:::|||||||||   ||||||          |||   :::::|||||||:::||
Db        180 eCysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGl 200

Qy        949 ACTTTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAAC 1008
              |:::::|||:::|||||||||||   ||||||   ||||||||||||||||||||||||
Db        200 nIleMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyTh 220

Qy       1009 ATCTACAGGAACACCTGATGAGGCTGCACCCTTGGCG---------------------- 1045
              ||||   ||||||||||||||||||||||||||||
Db        220 rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGl 240

Qy       1046 ----------------------------------------GAAAGAACACCTGACAC 1062
```

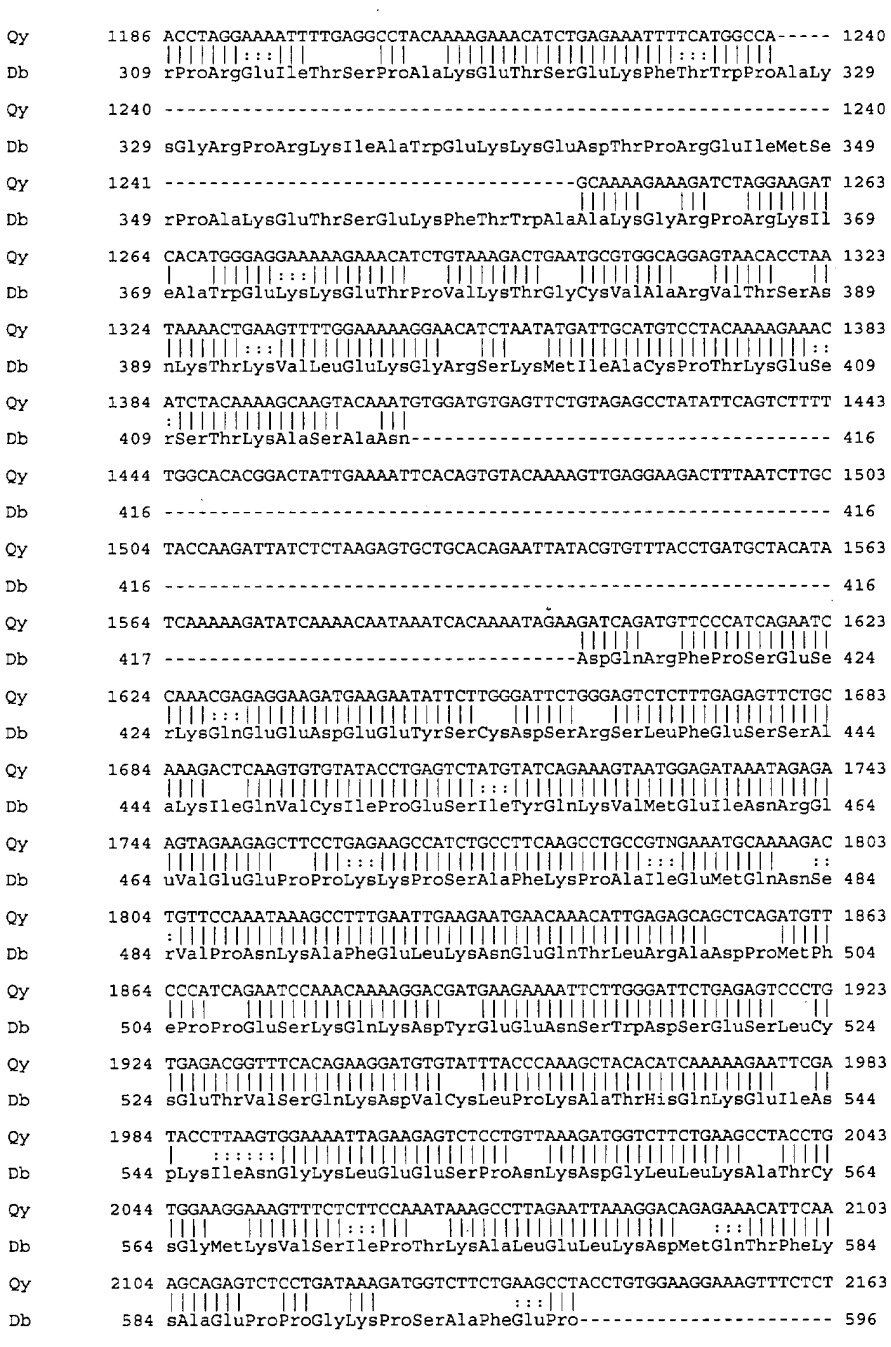


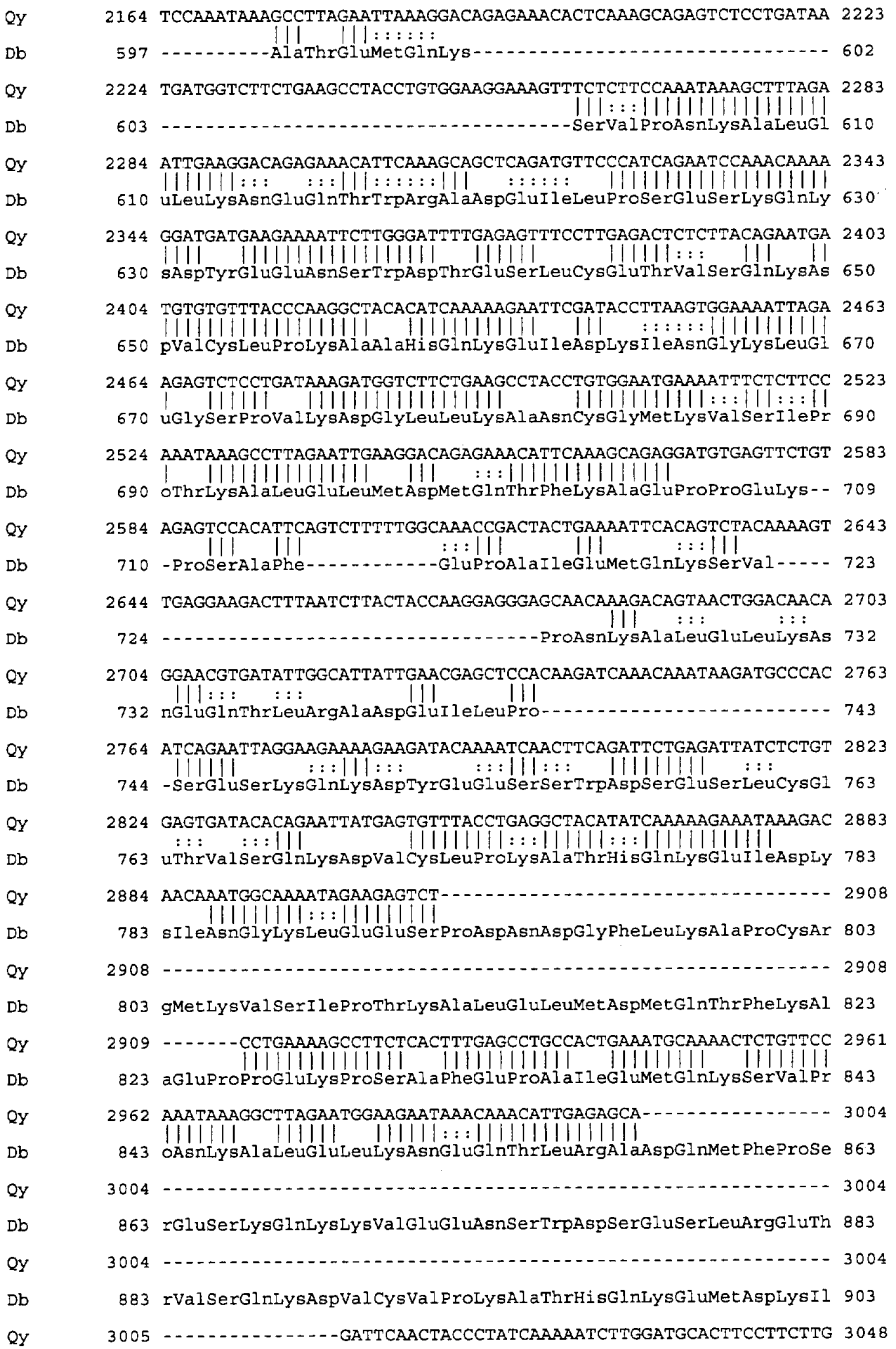


```
Qy       3004 ------------------------------------------------------------ 3004

Db        854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874

Qy       3004 ------------------------------------------------------------ 3004

Db        874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894

Qy       3005 ------------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTC 3045
                                |||||||||:::|||||||||||||||||||   :::   ||
Db        894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914

Qy       3046 TTGTGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGA 3105
              ||||||||||   |||||||||:::||||||:::|||||||||   |||   ||||||||
Db        914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934

Qy       3106 ACAAATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAA 3165
              ||||||||||   ||||||||||||||:::|||:::||||||||||||||||||||||||
Db        934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954

Qy       3166 ATCACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCC 3225
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974

Qy       3226 TTTAAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285
              ||||||||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGl 994

Qy       3286 C---------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGA 3330
                             |||||||||   ||||||:::||||||||||||   ||||||||
Db        994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014

Qy       3331 ACAGACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGT 3390
              ||||   ||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034

Qy       3391 TTCTCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGA 3450
              ||||||||||||||::||||||   |||   ||||||||||||||||||||||||||||
Db       1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

Qy       3451 AATTGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAA 3510
              ||||||||||||||||||||||:::||||||||||||||||||:::|||   ||||||||
Db       1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074

Qy       3511 TAAATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCT 3570
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094

Qy       3571 AAAACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGT 3630
              ||||||||||:::::::::::::|||||||||||||||||||||      ||||||||||
Db       1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114

Qy       3631 TCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              ||||   ||||||||||||||||||||||||||||||||||||
Db       1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128

RESULT 11
US-10-007-805-573
; Sequence 573, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 573
;   LENGTH: 1349
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-007-805-573

Alignment Scores:
Pred. No.:                3.27e-198      Length:          1349
Score:                    3440.00        Matches:         737
Percent Similarity:       65.39%         Conservative:    83
Best Local Similarity:    58.77%         Mismatches:      161
Query Match:              53.35%         Indels:          273
DB:                       13             Gaps:            12

US-09-602-362E-26 (1-3673) x US-10-007-805-573 (1-1349)

Qy        352 ACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTA 411
              |||      ||||||   :::|||||||||   :::|||   :::|||||||||||||||
Db         10 ThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAlaLeu 29

Qy        412 CACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGC 471
              ||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db         30 HisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLysCys 49

Qy        472 CNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAG 531
                 :::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         50 GlnProAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCysHis 69

Qy        532 AGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGT 591
              :::|||||| ||||||||||||||||||:::|||||||:::|||   |||||||||
Db         70 GlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAspVa 89

Qy        592 GTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAAC 651
              |||||||||   ||||||||||||||||||||||||||||   |||   :::|||||||
Db         89 lTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAlaLy 109

Qy        652 ACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTT 711
              |||||||||:::|||||||||||||||||   |||||||||||||||||||||||||||
Db        109 sLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLeuLe 129

Qy        712 ACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAA 771
              ||||:::|||   ||||||||||:::|||   |||||||||||||   ||||||||||||
Db        129 uLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAlaAs 149

Qy        772 TGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTC 831
              |||||||||   |||:::   |||||||||||||||||||||||:::|||   |||
Db        149 nAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGlyLe 169

Qy        832 ATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACA 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        169 uSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIleCy 189

Qy        892 TGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACT 951
              |||:::|||||||||   ||||||      |||   ::::::||||||:::|||::
Db        189 sGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGlnIl 209

Qy        952 TTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATC 1011
              ::::||:::|||||||||||||   ||||||   ||||||||||||||||||||||||
Db        209 eMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyThrSe 229

Qy       1012 TACAGGAACACCTGATGAGGCTGCACCCTTGGCG------------------------- 1045
              |   ||||||||||||||||||||||||||||||
Db        229 rAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGluSe 249

Qy       1046 -------------------------------------GAAAGAACACCTGACACGGC 1065
                                                    ||||||||||||||||||||
```


Db 249 rLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspThrAl 269
Qy 1066 TGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGC 1125
Db 269 aGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSerAs 289
Qy 1126 CAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAAC 1185
Db 289 pLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGluTh 309
Qy 1186 ACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA----- 1240
Db 309 rProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAlaLy 329
Qy 1240 ---------- 1240
Db 329 sGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMetSe 349
Qy 1241 ----------GCAAAAGAAAGATCTAGGAAGAT 1263
Db 349 rProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLysIl 369
Qy 1264 CACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAA 1323
Db 369 eAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSerAs 389
Qy 1324 TAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAAC 1383
Db 389 nLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGluSe 409
Qy 1384 ATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTT 1443
Db 409 rSerThrLysAlaSerAlaAsn---------- 416
Qy 1444 TGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGC 1503
Db 416 ---------- 416
Qy 1504 TACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACATA 1563
Db 416 ---------- 416
Qy 1564 TCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAATC 1623
Db 417 ----------AspGlnArgPheProSerGluSe 424
Qy 1624 CAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGC 1683
Db 424 rLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSerAl 444
Qy 1684 AAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGA 1743
Db 444 aLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGl 464
Qy 1744 AGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGAC 1803
Db 464 uValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSe 484
Qy 1804 TGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATGTT 1863
Db 484 rValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMetPh 504
Qy 1864 CCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCCTG 1923
Db 504 eProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLeuCy 524
Qy 1924 TGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGA 1983
Db 524 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 544
Qy 1984 TACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTG 2043
Db 544 pLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrCy 564
Qy 2044 TGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAA 2103
Db 564 sGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPheLy 584
Qy 2104 AGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCT 2163
Db 584 sAlaGluProProGlyLysProSerAlaPheGluPro---------- 596
Qy 2164 TCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAA 2223
Db 597 ----------AlaThrGluMetGlnLys---------- 602
Qy 2224 TGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGA 2283
Db 603 ----------SerValProAsnLysAlaLeuGl 610
Qy 2284 ATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAA 2343
Db 610 uLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysGlnLy 630
Qy 2344 GGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGA 2403
Db 630 sAspTyrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLysAs 650
Qy 2404 TGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGA 2463
Db 650 pValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGl 670
Qy 2464 AGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCC 2523
Db 670 uGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePr 690
Qy 2524 AAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGT 2583
Db 690 oThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys-- 709
Qy 2584 AGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGT 2643
Db 710 -ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal----- 723
Qy 2644 TGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACA 2703
Db 724 ----------ProAsnLysAlaLeuGluLeuLysAs 732
Qy 2704 GGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCAC 2763
Db 732 nGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 743
Qy 2764 ATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGT 2823
Db 744 -SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGl 763
Qy 2824 GAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGAC 2883
Db 763 uThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAspLy 783
Qy 2884 AACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 783 sIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCysAr 803
Qy 2908 ---------- 2908
Db 803 gMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAl 823
Qy 2909 -------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCC 2961
Db 823 aGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerValPr 843
Qy 2962 AAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 843 oAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPheProSe 863

```
Qy       3004 ------------------------------------------------------------ 3004

Db        863 rGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGluTh 883

Qy       3004 ------------------------------------------------------------ 3004

Db        883 rValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLysIl 903

Qy       3005 ----------------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTG 3048
                             ||||||||||:::|||||||||||||||||||   :::   |||||
Db        903 eSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSerCy 923

Qy       3049 TGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACA 3108
              |||||||   |||||||||:::|||||:::||||||||   |||   ||||||||||||
Db        923 sGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGluGl 943

Qy       3109 AATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATC 3168
              |||||||   |||||||||||||||:::|||:::|||||||||||||||||||||||||||
Db        943 nMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLysSe 963

Qy       3169 ACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTT 3228
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||   ||
Db        963 rGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuThrLe 983

Qy       3229 AAATCAAGAAGAAGAGAAGAAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC-- 3286
              ||||||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        983 uAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGluGl 1003

Qy       3287 ------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACA 3333
                         |||||||||   ||||||:::||||||||||||   ||||||||||||
Db       1003 uLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGluGl 1023

Qy       3334 GACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTC 3393
              |   ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       1023 nAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnValSe 1043

Qy       3394 TCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAAT 3453
              ||||||||||||:::||||||   |||   ||||||||||||||||||||||||||||||
Db       1043 rHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIl 1063

Qy       3454 TGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAA 3513
              ||||||||||||||||||:::|||||||||||||||||||:::|||   ||||||||||
Db       1063 eAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsnLy 1083

Qy       3514 ATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAA 3573
              ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db       1083 sTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLy 1103

Qy       3574 ACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCT 3633
              |||||||:::::::::::|||||||||||||||||||||      ||||||||||||||
Db       1103 sLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysValLe 1123

Qy       3634 GACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              |   |||||||||||||||||||||||||||||||||||||
Db       1123 uIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1136

RESULT 12
US-10-076-622-573
; Sequence 573, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C11
;  CURRENT APPLICATION NUMBER: US/10/076,622
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 573
;   LENGTH: 1349
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-076-622-573

Alignment Scores:
Pred. No.:                3.27e-198      Length:          1349
Score:                    3440.00        Matches:         737
Percent Similarity:       65.39%         Conservative:    83
Best Local Similarity:    58.77%         Mismatches:      161
Query Match:              53.35%         Indels:          273
DB:                       14             Gaps:            12

US-09-602-362E-26 (1-3673) x US-10-076-622-573 (1-1349)

Qy        352 ACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTA 411
              |||       ||||||   :::|||||||||   :::|||   :::|||||||||||||||||
Db         10 ThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAlaLeu 29

Qy        412 CACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGC 471
              ||||||||||||||||||||||      |||||||||||||||||||||||||||||||
Db         30 HisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLysCys 49

Qy        472 CNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAG 531
                 :::|||||||||||||||   ||||||||||||||||||||||||||||||||||
Db         50 GlnProAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCysHis 69

Qy        532 AGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGT 591
              :::|||||| ||||||||||||||||||:::|||||||:::|||   ||||||||
Db         70 GlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAspVa 89

Qy        592 GTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAAC 651
              ||||||||||   ||||||||||||||||||||||||||   |||   :::|||||||
Db         89 lTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAlaLy 109

Qy        652 ACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTT 711
              |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db        109 sLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLeuLe 129

Qy        712 ACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAA 771
              ||||:::|||   |||||||||:::|||   ||||||||||||||   ||||||||||
Db        129 uLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAlaAs 149

Qy        772 TGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTC 831
              |||||||||   |||:::   |||||||||||||||||||||||:::|||   |||
Db        149 nAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGlyLe 169

Qy        832 ATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACA 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 uSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIleCy 189

Qy        892 TGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACT 951
              |||:::|||||||||   ||||||          |||   ::::::||||||:::|||::
Db        189 sGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGlnIl 209

Qy        952 TTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATC 1011
              ::::|||:::|||||||||||   ||||||   ||||||||||||||||||||||||||
Db        209 eMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyThrSe 229

Qy       1012 TACAGGAACACCTGATGAGGCTGCACCCTTGGCG------------------------- 1045
              |   ||||||||||||||||||||||||||||
Db        229 rAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGluSe 249

Qy       1046 ----------------------------------------GAAAGAACACCTGACACGGC 1065
                                                     ||||||||||||||||||||
Db        249 rLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspThrAl 269

Qy       1066 TGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGC 1125
              ||||||||||:::||||||||||||||||||||||||   ||||||||||||||||||
Db        269 aGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSerAs 289

Qy       1126 CAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAAC 1185
              ||||||||||||||   ||||||||||||||||||||||||||||||   ||||||||
Db        289 pLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGluTh 309
```


Qy 1186 ACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA----- 1240
Db 309 rProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAlaLy 329
Qy 1240 ---------- 1240
Db 329 sGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMetSe 349
Qy 1241 ----------GCAAAAGAAAGATCTAGGAAGAT 1263
Db 349 rProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLysIl 369
Qy 1264 CACATGGGAGGAAAAAGAAACATCTGTAAAGACTGAATGCGTGGCAGGAGTAACACCTAA 1323
Db 369 eAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSerAs 389
Qy 1324 TAAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTGCATGTCCTACAAAAGAAAC 1383
Db 389 nLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGluSe 409
Qy 1384 ATCTACAAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGTCTTTT 1443
Db 409 rSerThrLysAlaSerAlaAsn---------- 416
Qy 1444 TGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGC 1503
Db 416 ---------- 416
Qy 1504 TACCAAGATTATCTCTAAGAGTGCTGCACAGAATTATACGTGTTTACCTGATGCTACATA 1563
Db 416 ---------- 416
Qy 1564 TCAAAAAGATATCAAAACAATAAATCACAAAATAGAAGATCAGATGTTCCCATCAGAATC 1623
Db 417 ----------AspGlnArgPheProSerGluSe 424
Qy 1624 CAAACGAGAGGAAGATGAAGAATATTCTTGGGATTCTGGGAGTCTCTTTGAGAGTTCTGC 1683
Db 424 rLysGlnGluGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSerAl 444
Qy 1684 AAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAGAGA 1743
Db 444 aLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArgGl 464
Qy 1744 AGTAGAAGAGCTTCCTGAGAAGCCATCTGCCTTCAAGCCTGCCGTNGAAATGCAAAAGAC 1803
Db 464 uValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSe 484
Qy 1804 TGTTCCAAATAAAGCCTTTGAATTGAAGAATGAACAAACATTGAGAGCAGCTCAGATGTT 1863
Db 484 rValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMetPh 504
Qy 1864 CCCATCAGAATCCAAACAAAAGGACGATGAAGAAAATTCTTGGGATTCTGAGAGTCCCTG 1923
Db 504 eProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLeuCy 524
Qy 1924 TGAGACGGTTTCACAGAAGGATGTGTATTTACCCAAAGCTACACATCAAAAAGAATTCGA 1983
Db 524 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 544
Qy 1984 TACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTTCTGAAGCCTACCTG 2043
Db 544 pLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrCy 564
Qy 2044 TGGAAGGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACATTCAA 2103
Db 564 sGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPheLy 584
Qy 2104 AGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCT 2163
Db 584 sAlaGluProProGlyLysProSerAlaPheGluPro---------- 596


Qy 2164 TCCAAATAAAGCCTTAGAATTAAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGATAA 2223
Db 597 ----------AlaThrGluMetGlnLys---------- 602
Qy 2224 TGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTTAGA 2283
Db 603 ----------SerValProAsnLysAlaLeuGl 610
Qy 2284 ATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAATCCAAACAAAA 2343
Db 610 uLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysGlnLy 630
Qy 2344 GGATGATGAAGAAAATTCTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAGAATGA 2403
Db 630 sAspTyrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLysAs 650
Qy 2404 TGTGTGTTTACCCAAGGCTACACATCAAAAAGAATTCGATACCTTAAGTGGAAAATTAGA 2463
Db 650 pValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGl 670
Qy 2464 AGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTTCC 2523
Db 670 uGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePr 690
Qy 2524 AAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTCTGT 2583
Db 690 oThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys-- 709
Qy 2584 AGAGTCCACATTCAGTCTTTTTGGCAAACCGACTACTGAAAATTCACAGTCTACAAAAGT 2643
Db 710 -ProSerAlaPhe----------GluProAlaIleGluMetGlnLysSerVal----- 723
Qy 2644 TGAGGAAGACTTTAATCTTACTACCAAGGAGGGAGCAACAAAGACAGTAACTGGACAACA 2703
Db 724 ----------ProAsnLysAlaLeuGluLeuLysAs 732
Qy 2704 GGAACGTGATATTGGCATTATTGAACGAGCTCCACAAGATCAAACAAATAAGATGCCCAC 2763
Db 732 nGluGlnThrLeuArgAlaAspGluIleLeuPro---------- 743
Qy 2764 ATCAGAATTAGGAAGAAAAGAAGATACAAAATCAACTTCAGATTCTGAGATTATCTCTGT 2823
Db 744 -SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCysGl 763
Qy 2824 GAGTGATACACAGAATTATGAGTGTTTACCTGAGGCTACATATCAAAAAGAAATAAAGAC 2883
Db 763 uThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAspLy 783
Qy 2884 AACAAATGGCAAAATAGAAGAGTCT---------- 2908
Db 783 sIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCysAr 803
Qy 2908 ---------- 2908
Db 803 gMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAl 823
Qy 2909 -------CCTGAAAAGCCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTTCC 2961
Db 823 aGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerValPr 843
Qy 2962 AAATAAAGGCTTAGAATGGAAGAATAAACAAACATTGAGAGCA---------- 3004
Db 843 oAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPheProSe 863
Qy 3004 ---------- 3004
Db 863 rGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGluTh 883
Qy 3004 ---------- 3004
Db 883 rValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLysIl 903
Qy 3005 ----------GATTCAACTACCCTATCAAAAATCTTGGATGCACTTCCTTCTTG 3048

```
                          |||||||||:::|||||||||||||||||||   :::   |||||
Db        903 eSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSerCy 923

Qy       3049 TGAAAGAGGAAGGGAACTTAAAAAAGATAACTGTGAACAAATTACAGCAAAAATGGAACA 3108
              |||||||   |||||||||:::||||||:::|||||||||   |||   ||||||||||
Db        923 sGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGluGl 943

Qy       3109 AATGAAAAATAAGTTTTGTGTACTACAAAAGGAACTGTCAGAAGCGAAAGAAATAAAATC 3168
              |||||||   |||||||||||||||:::|||:::|||||||||||||||||||||||||||
Db        943 nMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLysSe 963

Qy       3169 ACAGTTAGAGAACCAAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTGCCTTT 3228
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||   ||
Db        963 rGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuThrLe 983

Qy       3229 AAATCAAGAAGAAGAGAAGAGAAGAAATGTCGATATATTAAAAGAAAAAATTAGACCC-- 3286
              |||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        983 uAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgGluGl 1003

Qy       3287 ------------GAAGAGCAACTTAGGAAAAAGTTAGAAGTGAAACACCAACTTGAACA 3333
                         |||||||||   ||||||:::|||||||||||   ||||||||||
Db       1003 uLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGluGl 1023

Qy       3334 GACTCTCAGAATACAAGATATAGAATTGAAAAGTGTAACAAGTAATTTGAATCAGGTTTC 3393
              |   |||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       1023 nAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnValSe 1043

Qy       3394 TCACACTCATGAAAGTGAAAATGATCTCTTTCATGAAAATTGCATGTTGAAAAAGGAAAT 3453
              |||||||||||||:::||||||   |||   ||||||||||||||||||||||||||||
Db       1043 rHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIl 1063

Qy       3454 TGCCATGCTAAAACTGGAAGTAGCCACACTGAAACATCAACACCAGGTGAAGGAAAATAA 3513
              |||||||||||||||||||:::|||||||||||||||||:::|||   ||||||||||
Db       1063 eAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsnLy 1083

Qy       3514 ATACTTTGAGGACATTAAGATTTTACAAGAAAAGAATGCTGAACTTCAAATGACCCTAAA 3573
              |||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db       1083 sTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLy 1103

Qy       3574 ACTGAAACAGAAAACAGTAACAAAAAGGGCATCTCAGTATAGAGAGCAGCTTAAAGTTCT 3633
              |||||||:::::::::::|||||||||||||||||||||     ||||||||||||||
Db       1103 sLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysValLe 1123

Qy       3634 GACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
              |   ||||||||||||||||||||||||||||||||||||
Db       1123 uIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1136

RESULT 13
US-10-124-805-573
; Sequence 573, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C12
;  CURRENT APPLICATION NUMBER: US/10/124,805
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 573
;   LENGTH: 1349
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-124-805-573

Alignment Scores:
Pred. No.:              3.27e-198       Length:         1349
Score:                  3440.00         Matches:        737
Percent Similarity:     65.39%          Conservative:   83
Best Local Similarity:  58.77%          Mismatches:     161

Query Match:            53.35%          Indels:         273
DB:                     14              Gaps:           12

US-09-602-362E-26 (1-3673) x US-10-124-805-573 (1-1349)

Qy        352 ACAGTAGGGAAGAAGCCCGTCAACCTGAACAAAAGAGATATGAAGAAGAGGACTGCTCTA 411
              |||       ||||||   :::|||||||||   :::|||   :::|||||||||||||||
Db         10 ThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAlaLeu 29

Qy        412 CACTGGGCCTGTGTCAATGGCCATGCANAAGTAGTAACATTTCTGGTAGACAGAAAGTGC 471
              ||||||||||||||||||||||||      ||||||||||||||||||||||||||||||
Db         30 HisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLysCys 49

Qy        472 CNGCTTAATGTCCTTGATGGCGAAGGGAGGACACCTCTGATGAAGGCTCTACAATGCGAG 531
                 :::|||||||||||||||   |||||||||||||||||||||||||||||
Db         50 GlnProAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCysHis 69

Qy        532 AGGGAAGCTTTGTGCAAATATTCTCATAGATGCTGGTGCTGATCTAAATTATGTAGATGT 591
              :::|||||| |||||||||||||||||||:::|||||||:::|||   |||||||
Db         70 GlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAspVa 89

Qy        592 GTATGGCAACACGGCTCTCCATTATGCCGTTTATAGTGAGAATTTATTAATGGTGGCAAC 651
              |||||||||   ||||||||||||||||||||||||||||   |||   :::||||||
Db         89 lTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAlaLy 109

Qy        652 ACTGCTGTCCTATGGTGCAGTCATCGAGGTGCAAAACAAGGCTAGCCTCACACCCCTTTT 711
              |||||||||:::||||||||||||||||||||   ||||||||||||||||||||||||
Db        109 sLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLeuLe 129

Qy        712 ACTGGCCATACAGAAAAGAAGCAAGCAAACTGTGGAATTTTTACTAACAAAAAATGCAAA 771
              ||||:::|||   |||||||||:::|||   ||||||||||||||   ||||||||||
Db        129 uLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAlaAs 149

Qy        772 TGCAAACGCATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGCCATATGTGAAGGCTC 831
              |||||||||   |||:::   |||||||||||||||||||||||:::|||   |||
Db        149 nAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGlyLe 169

Qy        832 ATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGTCTTTGCTGAAGACATACA 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 uSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIleCy 189

Qy        892 TGGAATAACTGCAGAACGTTATGCTGCTGCTCGTGGAGTTAATTACATTCATCAACAACT 951
              |||:::|||||||||   ||||||         |||   ::::::||||||:::|||::
Db        189 sGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGlnIl 209

Qy        952 TTTGGAACATATACGAAAATTACCTAAAAATCCTCAAAATACCAATCCAGAAGGAACATC 1011
              ::::|||:::|||||||||||   ||||||   |||||||||||||||||||||||||||
Db        209 eMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyThrSe 229

Qy       1012 TACAGGAACACCTGATGAGGCTGCACCCTTGGCG---------------------------- 1045
              |   |||||||||||||||||||||||||||||
Db        229 rAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGluSe 249

Qy       1046 ----------------------------------------GAAAGAACACCTGACACGGC 1065
                                                     ||||||||||||||||||||
Db        249 rLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspThrAl 269

Qy       1066 TGAAAGCTTGCTGGAAAAAACACCTGACGAGGCTGCACGCTTGGTGGAGGGAACGTCTGC 1125
              ||||||||||:::|||||||||||||||||||||||||   ||||||||||||||||||
Db        269 aGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSerAs 289

Qy       1126 CAAAATTCAATGTCTGGGGAAAGCAACATCTGGAAAGTTTGAACAGTCAACAGAAGAAAC 1185
              ||||||||||||||   |||||||||||||||||||||||||||||   ||||||||
Db        289 pLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGluTh 309

Qy       1186 ACCTAGGAAAATTTTGAGGCCTACAAAAGAAACATCTGAGAAATTTTCATGGCCA----- 1240
              |||||||:::|||   |||   ||||||||||||||||||||:::||||||
Db        309 rProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAlaLy 329

Qy       1240 ------------------------------------------------------------ 1240

Db        329 sGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMetSe 349
```